

35th AAAI Conference on Artificial Intelligence (AAAI-21)

Volume 2: AAAI Technical Tracks

- Computer Vision I

Online
2 – 9 February 2021

Part 1 of 2

ISBN: 978-1-7138-3588-2

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2021) by Association for the Advancement of Artificial Intelligence
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

For permission requests, please contact Association for the Advancement of Artificial Intelligence
at the address below.

Association for the Advancement of Artificial Intelligence
2275 East Bayshore Road
Suite 160
Palo Alto, California 94303
USA

Phone: 1-650-328-3123
Fax: 1-650-321-4457

<https://aaai.org/Press/press.php>

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
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

PART 1

AAAI TECHNICAL TRACK ON COMPUTER VISION I

| | |
|---|-----|
| LOCALIZATION IN THE CROWD WITH TOPOLOGICAL CONSTRAINTS | 872 |
| <i>Shahira Abousamra, Minh Hoai, Dimitris Samaras, Chao Chen</i> | |
| DEEP EVENT STEREO LEVERAGED BY EVENT-TO-IMAGE TRANSLATION | 882 |
| <i>Soikat Hasan Ahmed, Hae Woong Jang, S M Nadim Uddin, Yong Ju Jung</i> | |
| OPTICAL FLOW ESTIMATION FROM A SINGLE MOTION-BLURRED IMAGE | 891 |
| <i>Dawit Mureja Argaw, Junsik Kim, Francois Rameau, Jae Won Cho, In So Kweon</i> | |
| MOTION-BLURRED VIDEO INTERPOLATION AND EXTRAPOLATION | 901 |
| <i>Dawit Mureja Argaw, Junsik Kim, Francois Rameau, In So Kweon</i> | |
| DISENTANGLED MULTI-RELATIONAL GRAPH CONVOLUTIONAL NETWORK FOR PEDESTRIAN TRAJECTORY PREDICTION | 911 |
| <i>Inhwan Bae, Hae-Gon Jeon</i> | |
| DENSE EVENTS GROUNDING IN VIDEO | 920 |
| <i>Peijun Bao, Qian Zheng, Yadong Mu</i> | |
| CONTEXT-AWARE ATTENTIONAL POOLING (CAP) FOR FINE-GRAINED VISUAL CLASSIFICATION | 929 |
| <i>Ardhendu Behera, Zachary Wharton, Pradeep R P G Hewage, Asish Bera</i> | |
| APPEARANCE-MOTION MEMORY CONSISTENCY NETWORK FOR VIDEO ANOMALY DETECTION | 938 |
| <i>Ruichu Cai, Hao Zhang, Wen Liu, Shenghua Gao, Zhifeng Hao</i> | |
| RETHINKING OBJECT DETECTION IN RETAIL STORES | 947 |
| <i>Yuanqiang Cai, Longyin Wen, Libo Zhang, Dawei Du, Weiqiang Wang</i> | |
| YOLOBILE: REAL-TIME OBJECT DETECTION ON MOBILE DEVICES VIA COMPRESSION-COMPILATION CO-DESIGN | 955 |
| <i>Yuxuan Cai, Hongjia Li, Geng Yuan, Wei Niu, Yanyu Li, Xulong Tang, Bin Ren, Yanzhi Wang</i> | |
| SEMANTIC MAPNET: BUILDING ALLOCENTRIC SEMANTIC MAPS AND REPRESENTATIONS FROM EGOCENTRIC VIEWS | 964 |
| <i>Vincent Cartillier, Zhile Ren, Neha Jain, Stefan Lee, Irfan Essa, Dhruv Batra</i> | |
| UNDERSTANDING DEFORMABLE ALIGNMENT IN VIDEO SUPER-RESOLUTION | 973 |
| <i>Kelvin C. K. Chan, Xintao Wang, Ke Yu, Chao Dong, Chen Change Loy</i> | |
| DEEP METRIC LEARNING WITH GRAPH CONSISTENCY | 982 |
| <i>Binghui Chen, Pengyu Li, Zhaoyi Yan, Biao Wang, Lei Zhang</i> | |
| CNN PROFILER ON POLAR COORDINATE IMAGES FOR TROPICAL CYCLONE STRUCTURE ANALYSIS | 991 |
| <i>Boyo Chen, Buo-Fu Chen, Chun Min Hsiao</i> | |

| | |
|--|------|
| COMMONSENSE KNOWLEDGE AWARE CONCEPT SELECTION FOR DIVERSE AND INFORMATIVE VISUAL STORYTELLING | 999 |
| <i>Hong Chen, Yifei Huang, Hiroya Takamura, Hideki Nakayama</i> | |
| ATTENTION-BASED MULTI-LEVEL FUSION NETWORK FOR LIGHT FIELD DEPTH ESTIMATION | 1009 |
| <i>Jiixin Chen, Shuo Zhang, Youfang Lin</i> | |
| JOINT DEMOSAICKING AND DENOISING IN THE WILD: THE CASE OF TRAINING UNDER GROUND TRUTH UNCERTAINTY | 1018 |
| <i>Jierun Chen, Song Wen, S.-H. Gary Chan</i> | |
| SPATIAL-TEMPORAL CAUSAL INFERENCE FOR PARTIAL IMAGE-TO-VIDEO ADAPTATION | 1027 |
| <i>Jin Chen, Xinxiao Wu, Yao Hu, Jiebo Luo</i> | |
| REF-NMS: BREAKING PROPOSAL BOTTLENECKS IN TWO-STAGE REFERRING EXPRESSION GROUNDING | 1036 |
| <i>Long Chen, Wenbo Ma, Jun Xiao, Hanwang Zhang, Shih-Fu Chang</i> | |
| RSPNET: RELATIVE SPEED PERCEPTION FOR UNSUPERVISED VIDEO REPRESENTATION LEARNING | 1045 |
| <i>Peihao Chen, Deng Huang, Dongliang He, Xiang Long, Runhao Zeng, Shilei Wen, Mingkui Tan, Chuang Gan</i> | |
| DUAL DISTRIBUTION ALIGNMENT NETWORK FOR GENERALIZABLE PERSON RE-IDENTIFICATION | 1054 |
| <i>Peixian Chen, Pingyang Dai, Jianzhuang Liu, Feng Zheng, Mingliang Xu, Qi Tian, Rongrong Ji</i> | |
| RGB-D SALIENT OBJECT DETECTION VIA 3D CONVOLUTIONAL NEURAL NETWORKS | 1063 |
| <i>Qian Chen, Ze Liu, Yi Zhang, Keren Fu, Qijun Zhao, Hongwei Du</i> | |
| MIND-THE-GAP! UNSUPERVISED DOMAIN ADAPTATION FOR TEXT-VIDEO RETRIEVAL | 1072 |
| <i>Qingchao Chen, Yang Liu, Samuel Albanie</i> | |
| LOCAL RELATION LEARNING FOR FACE FORGERY DETECTION | 1081 |
| <i>Shen Chen, Taiping Yao, Yang Chen, Shouhong Ding, Jilin Li, Rongrong Ji</i> | |
| DEDUCTIVE LEARNING FOR WEAKLY-SUPERVISED 3D HUMAN POSE ESTIMATION VIA UNCALIBRATED CAMERAS | 1089 |
| <i>Xipeng Chen, Pengxu Wei, Liang Lin</i> | |
| A UNIFIED MULTI-SCENARIO ATTACKING NETWORK FOR VISUAL OBJECT TRACKING | 1097 |
| <i>Xuesong Chen, Canmiao Fu, Feng Zheng, Yong Zhao, Hongsheng Li, Ping Luo, Guo-Jun Qi</i> | |
| SSD-GAN: MEASURING THE REALNESS IN THE SPATIAL AND SPECTRAL DOMAINS | 1105 |
| <i>Yuanqi Chen, Ge Li, Cece Jin, Shan Liu, Thomas Li</i> | |
| MULTI-SCALE SPATIAL TEMPORAL GRAPH CONVOLUTIONAL NETWORK FOR SKELETON-BASED ACTION RECOGNITION | 1113 |
| <i>Zhan Chen, Sicheng Li, Bing Yang, Qinghan Li, Hong Liu</i> | |

| | |
|---|------|
| CASCADE NETWORK WITH GUIDED LOSS AND HYBRID ATTENTION FOR FINDING GOOD CORRESPONDENCES | 1123 |
| <i>Zhi Chen, Fan Yang, Wenbing Tao</i> | |
| GENERALIZABLE REPRESENTATION LEARNING FOR MIXTURE DOMAIN FACE ANTI-SPOOFING | 1132 |
| <i>Zhihong Chen, Taiping Yao, Kekai Sheng, Shouhong Ding, Ying Tai, Jilin Li, Feiyue Huang, Xinyu Jin</i> | |
| SSPC-NET: SEMI-SUPERVISED SEMANTIC 3D POINT CLOUD SEGMENTATION NETWORK | 1140 |
| <i>Mingmei Cheng, Le Hui, Jin Xie, Jian Yang</i> | |
| DEEP FEATURE SPACE TROJAN ATTACK OF NEURAL NETWORKS BY CONTROLLED DETOXIFICATION | 1148 |
| <i>Siyuan Cheng, Yingqi Liu, Shiqing Ma, Xiangyu Zhang</i> | |
| GRAPH AND TEMPORAL CONVOLUTIONAL NETWORKS FOR 3D MULTI-PERSON POSE ESTIMATION IN MONOCULAR VIDEOS | 1157 |
| <i>Yu Cheng, Bo Wang, Bo Yang, Robby T. Tan</i> | |
| DRAMAQA: CHARACTER-CENTERED VIDEO STORY UNDERSTANDING WITH HIERARCHICAL QA | 1166 |
| <i>Seongho Choi, Kyoung-Woon On, Yu-Jung Heo, Ahjeong Seo, Youwon Jang, Minsu Lee, Byoung-Tak Zhang</i> | |
| DEEPCOLLABORATION: COLLABORATIVE GENERATIVE AND DISCRIMINATIVE MODELS FOR CLASS INCREMENTAL LEARNING | 1175 |
| <i>Bo Cui, Guyue Hu, Shan Yu</i> | |
| SPLIT THEN REFINE: STACKED ATTENTION-GUIDED RESUNETS FOR BLIND SINGLE IMAGE VISIBLE WATERMARK REMOVAL | 1184 |
| <i>Xiaodong Cun, Chi-Man Pun</i> | |
| RSGNET: RELATION BASED SKELETON GRAPH NETWORK FOR CROWDED SCENES POSE ESTIMATION | 1193 |
| <i>Yan Dai, Xuanhan Wang, Lianli Gao, Jingkuan Song, Heng Tao Shen</i> | |
| VOXEL R-CNN: TOWARDS HIGH PERFORMANCE VOXEL-BASED 3D OBJECT DETECTION | 1201 |
| <i>Jiajun Deng, Shaoshuai Shi, Peiwei Li, Wengang Zhou, Yanyong Zhang, Houqiang Li</i> | |
| ARBITRARY VIDEO STYLE TRANSFER VIA MULTI-CHANNEL CORRELATION | 1210 |
| <i>Yingying Deng, Fan Tang, Weiming Dong, Haibin Huang, Chongyang Ma, Changsheng Xu</i> | |
| SIMILARITY REASONING AND FILTRATION FOR IMAGE-TEXT MATCHING | 1218 |
| <i>Haiwen Diao, Ying Zhang, Lin Ma, Huchuan Lu</i> | |
| SPATIO-TEMPORAL DIFFERENCE DESCRIPTOR FOR SKELETON-BASED ACTION RECOGNITION | 1227 |
| <i>Chongyang Ding, Kai Liu, Jari Korhonen, Evgeny Belyaev</i> | |
| TOWARDS UNIVERSAL PHYSICAL ATTACKS ON SINGLE OBJECT TRACKING | 1236 |
| <i>Li Ding, Yongwei Wang, Kaiwen Yuan, Minyang Jiang, Ping Wang, Hua Huang, Z. Jane Wang</i> | |

| | |
|--|------|
| MODELING THE PROBABILISTIC DISTRIBUTION OF UNLABELED DATA FOR ONE-SHOT MEDICAL IMAGE SEGMENTATION | 1246 |
| <i>Yuhang Ding, Xin Yu, Yi Yang</i> | |
| FEW-SHOT CLASS-INCREMENTAL LEARNING VIA RELATION KNOWLEDGE DISTILLATION | 1255 |
| <i>Songlin Dong, Xiaopeng Hong, Xiaoyu Tao, Xinyuan Chang, Xing Wei, Yihong Gong</i> | |
| MIEHDR CNN: MAIN IMAGE ENHANCEMENT BASED GHOST-FREE HIGH DYNAMIC RANGE IMAGING USING DUAL-LENS SYSTEMS | 1264 |
| <i>Xuan Dong, Xiaoyan Hu, Weixin Li, Xiaojie Wang, Yunhong Wang</i> | |
| BOOSTING IMAGE-BASED MUTUAL GAZE DETECTION USING PSEUDO 3D GAZE | 1273 |
| <i>Bardia Doosti, Ching-Hui Chen, Raviteja Vemulapalli, Xuhui Jia, Yukun Zhu, Bradley Green</i> | |
| HOW TO SAVE YOUR ANNOTATION COST FOR PANOPTIC SEGMENTATION? | 1282 |
| <i>Xuefeng Du, Chenhan Jiang, Hang Xu, Gengwei Zhang, Zhenguo Li</i> | |
| DIRV: DENSE INTERACTION REGION VOTING FOR END-TO-END HUMAN-OBJECT INTERACTION DETECTION | 1291 |
| <i>Hao-Shu Fang, Yichen Xie, Dian Shao, Cewu Lu</i> | |
| DECAUG: AUGMENTING HOI DETECTION VIA DECOMPOSITION | 1300 |
| <i>Hao-Shu Fang, Yichen Xie, Dian Shao, Yong-Lu Li, Cewu Lu</i> | |
| PARTIALLY NON-AUTOREGRESSIVE IMAGE CAPTIONING | 1309 |
| <i>Zhengcong Fei</i> | |
| MEMORY-AUGMENTED IMAGE CAPTIONING | 1317 |
| <i>Zhengcong Fei</i> | |
| EDGE-COMPETING PATHOLOGICAL LIVER VESSEL SEGMENTATION WITH LIMITED LABELS | 1325 |
| <i>Zunlei Feng, Zhonghua Wang, Xinchao Wang, Xiuming Zhang, Lechao Cheng, Jie Lei, Yuexuan Wang, Mingli Song</i> | |
| VISUAL BOUNDARY KNOWLEDGE TRANSLATION FOR FOREGROUND SEGMENTATION | 1334 |
| <i>Zunlei Feng, Lechao Cheng, Xinchao Wang, Xiang Wang, Ya Jie Liu, Xiangtong Du, Mingli Song</i> | |

PART 2

| | |
|--|------|
| LEARNING COMPLEX 3D HUMAN SELF-CONTACT | 1343 |
| <i>Mihai Fieraru, Mihai Zanfir, Elisabeta Oneata, Alin-Ionut Popa, Vlad Olaru, Cristian Sminchisescu</i> | |
| RAIN STREAK REMOVAL VIA DUAL GRAPH CONVOLUTIONAL NETWORK | 1352 |
| <i>Xueyang Fu, Qi Qi, Zheng-Jun Zha, Yurui Zhu, Xinghao Ding</i> | |
| COMPFEAT: COMPREHENSIVE FEATURE AGGREGATION FOR VIDEO INSTANCE SEGMENTATION | 1361 |
| <i>Yang Fu, Linjie Yang, Ding Liu, Thomas S. Huang, Humphrey Shi</i> | |
| DEEP METRIC LEARNING WITH SELF-SUPERVISED RANKING | 1370 |
| <i>Zheren Fu, Yan Li, Zhendong Mao, Quan Wang, Yongdong Zhang</i> | |

| | |
|--|------|
| A SYSTEMATIC EVALUATION OF OBJECT DETECTION NETWORKS FOR SCIENTIFIC PLOTS | 1379 |
| <i>Pritha Ganguly, Nitesh S Methani, Mitesh M. Khapra, Pratyush Kumar</i> | |
| THE COMPLEXITY OF OBJECT ASSOCIATION IN MULTIPLE OBJECT TRACKING | 1388 |
| <i>Robert Galian, Thekla Hamm, Sebastian Ordyniak</i> | |
| LEARNING LOCAL NEIGHBORING STRUCTURE FOR ROBUST 3D SHAPE REPRESENTATION | 1397 |
| <i>Zhongpai Gao, Junchi Yan, Guangtao Zhai, Juyong Zhang, Yiyang Yang, Xiaokang Yang</i> | |
| SEMANTIC-GUIDED REINFORCED REGION EMBEDDING FOR GENERALIZED ZERO-SHOT LEARNING | 1406 |
| <i>Jiannan Ge, Hongtao Xie, Shaobo Min, Yongdong Zhang</i> | |
| DYNAMIC GRAPH REPRESENTATION LEARNING FOR VIDEO DIALOG VIA MULTI-MODAL SHUFFLED TRANSFORMERS | 1415 |
| <i>Shijie Geng, Peng Gao, Moitreyia Chatterjee, Chiori Hori, Jonathan Le Roux, Yongfeng Zhang, Hongsheng Li, Anoop Cherian</i> | |
| BOUNDARY-AWARE GEOMETRIC ENCODING FOR SEMANTIC SEGMENTATION OF POINT CLOUDS | 1424 |
| <i>Jingyu Gong, Jiachen Xu, Xin Tan, Jie Zhou, Yanyun Qu, Yuan Xie, Lizhuang Ma</i> | |
| ANALOGICAL IMAGE TRANSLATION FOR FOG GENERATION | 1433 |
| <i>Rui Gong, Dengxin Dai, Yuhua Chen, Wen Li, Danda Pani Paudel, Luc Van Gool</i> | |
| TEMPORAL ROI ALIGN FOR VIDEO OBJECT RECOGNITION | 1442 |
| <i>Tao Gong, Kai Chen, Xinjiang Wang, Qi Chu, Feng Zhu, Dahua Lin, Nenghai Yu, Huamin Feng</i> | |
| SMART FRAME SELECTION FOR ACTION RECOGNITION | 1451 |
| <i>Shreyank N Gowda, Marcus Rohrbach, Laura Sevilla-Lara</i> | |
| PROXY SYNTHESIS: LEARNING WITH SYNTHETIC CLASSES FOR DEEP METRIC LEARNING | 1460 |
| <i>Geonmo Gu, Byungsoo Ko, Han-Gyu Kim</i> | |
| INTERPRETABLE GRAPH CAPSULE NETWORKS FOR OBJECT RECOGNITION | 1469 |
| <i>Jindong Gu</i> | |
| CLASS-INCREMENTAL INSTANCE SEGMENTATION VIA MULTI-TEACHER NETWORKS | 1478 |
| <i>Yanan Gu, Cheng Deng, Kun Wei</i> | |
| EFFICIENTDERAIN: LEARNING PIXEL-WISE DILATION FILTERING FOR HIGH-EFFICIENCY SINGLE-IMAGE DERAINING | 1487 |
| <i>Qing Guo, Jingyang Sun, Felix Juefei-Xu, Lei Ma, Xiaofei Xie, Wei Feng, Yang Liu, Jianjun Zhao</i> | |
| ORDER REGULARIZATION ON ORDINAL LOSS FOR HEAD POSE, AGE AND GAZE ESTIMATION | 1496 |
| <i>Tianchu Guo, Hui Zhang, Byungin Yoo, Yongchao Liu, Youngjun Kwak, Jae-Joon Han</i> | |
| DECOUPLED AND MEMORY-REINFORCED NETWORKS: TOWARDS EFFECTIVE FEATURE LEARNING FOR ONE-STEP PERSON SEARCH | 1505 |
| <i>Chuchu Han, Zhedong Zheng, Changxin Gao, Nong Sang, Yi Yang</i> | |

| | |
|---|------|
| SPHERICAL IMAGE GENERATION FROM A SINGLE IMAGE BY CONSIDERING SCENE SYMMETRY | 1513 |
| <i>Takayuki Hara, Yusuke Mukuta, Tatsuya Harada</i> | |
| PROGRESSIVE ONE-SHOT HUMAN PARSING | 1522 |
| <i>Haoyu He, Jing Zhang, Bhavani Thuraisingham, Dacheng Tao</i> | |
| CONSISTENT-SEPARABLE FEATURE REPRESENTATION FOR SEMANTIC SEGMENTATION | 1531 |
| <i>Xingjian He, Jing Liu, Jun Fu, Xinxin Zhu, Jinqiao Wang, Hanqing Lu</i> | |
| ERROR-AWARE DENSITY ISOMORPHISM RECONSTRUCTION FOR UNSUPERVISED CROSS-DOMAIN CROWD COUNTING | 1540 |
| <i>Yuhang He, Zhiheng Ma, Xing Wei, Xiaopeng Hong, Wei Ke, Yihong Gong</i> | |
| DROPLOSS FOR LONG-TAIL INSTANCE SEGMENTATION | 1549 |
| <i>Ting-I Hsieh, Esther Robb, Hwann-Tzong Chen, Jia-Bin Huang</i> | |
| HAND-MODEL-AWARE SIGN LANGUAGE RECOGNITION | 1558 |
| <i>Hezhen Hu, Wengang Zhou, Houqiang Li</i> | |
| STRATIFIED RULE-AWARE NETWORK FOR ABSTRACT VISUAL REASONING | 1567 |
| <i>Sheng Hu, Yuqing Ma, Xianglong Liu, Yanlu Wei, Shihao Bai</i> | |
| VIVO: VISUAL VOCABULARY PRE-TRAINING FOR NOVEL OBJECT CAPTIONING | 1575 |
| <i>Xiaowei Hu, Xi Yin, Kevin Lin, Lei Zhang, Jianfeng Gao, Lijuan Wang, Zicheng Liu</i> | |
| EXPLOITING RELATIONSHIP FOR COMPLEX-SCENE IMAGE GENERATION | 1584 |
| <i>Tianyu Hua, Hongdong Zheng, Yalong Bai, Wei Zhang, Xiao-Ping Zhang, Tao Mei</i> | |
| MODELING DEEP LEARNING BASED PRIVACY ATTACKS ON PHYSICAL MAIL | 1593 |
| <i>Bingyao Huang, Ruyi Lian, Dimitris Samaras, Haibin Ling</i> | |
| PTN: A POISSON TRANSFER NETWORK FOR SEMI-SUPERVISED FEW-SHOT LEARNING | 1602 |
| <i>Huaxi Huang, Junjie Zhang, Jian Zhang, Qiang Wu, Chang Xu</i> | |
| TEXT-GUIDED GRAPH NEURAL NETWORKS FOR REFERRING 3D INSTANCE SEGMENTATION | 1610 |
| <i>Pin-Hao Huang, Han-Hung Lee, Hwann-Tzong Chen, Tyng-Luh Liu</i> | |
| INITIATIVE DEFENSE AGAINST FACIAL MANIPULATION | 1619 |
| <i>Qidong Huang, Jie Zhang, Wenbo Zhou, Weiming Zhang, Nenghai Yu</i> | |
| SNAPMIX: SEMANTICALLY PROPORTIONAL MIXING FOR AUGMENTING FINE-GRAINED DATA | 1628 |
| <i>Shaoli Huang, Xinchao Wang, Dacheng Tao</i> | |
| A HYBRID ATTENTION MECHANISM FOR WEAKLY-SUPERVISED TEMPORAL ACTION LOCALIZATION | 1637 |
| <i>Ashrafal Islam, Chengjiang Long, Richard Radke</i> | |
| CONTEXT-AWARE GRAPH CONVOLUTION NETWORK FOR TARGET RE-IDENTIFICATION | 1646 |
| <i>Deyi Ji, Haoran Wang, Hanzhe Hu, Weihao Gan, Wei Wu, Junjie Yan</i> | |

| | |
|--|------|
| IMPROVING IMAGE CAPTIONING BY LEVERAGING INTRA- AND INTER-LAYER GLOBAL REPRESENTATION IN TRANSFORMER NETWORK | 1655 |
| <i>Jiayi Ji, Yunpeng Luo, Xiaoshuai Sun, Fuhai Chen, Gen Luo, Yongjian Wu, Yue Gao, Rongrong Ji</i> | |
| FREQUENCY CONSISTENT ADAPTATION FOR REAL WORLD SUPER RESOLUTION | 1664 |
| <i>Xiaozhong Ji, Guangpin Tao, Yun Cao, Ying Tai, Tong Lu, Chengjie Wang, Jilin Li, Feiyue Huang</i> | |
| MATCHING ON SETS: CONQUER OCCLUDED PERSON RE-IDENTIFICATION WITHOUT ALIGNMENT | 1673 |
| <i>Mengxi Jia, Xinhua Cheng, Yunpeng Zhai, Shijian Lu, Siwei Ma, Yonghong Tian, Jian Zhang</i> | |
| GRADINGNET: TOWARDS PROVIDING RELIABLE SUPERVISIONS FOR WEAKLY SUPERVISED OBJECT DETECTION BY GRADING THE BOX CANDIDATES | 1682 |
| <i>Qifei Jia, Shikui Wei, Tao Ruan, Yufeng Zhao, Yao Zhao</i> | |
| SSN3D: SELF-SEPARATED NETWORK TO ALIGN PARTS FOR 3D CONVOLUTION IN VIDEO PERSON RE-IDENTIFICATION | 1691 |
| <i>Xiaoke Jiang, Yu Qiao, Junjie Yan, Qichen Li, Wanrong Zheng, Dapeng Chen</i> | |
| TRAINING BINARY NEURAL NETWORK WITHOUT BATCH NORMALIZATION FOR IMAGE SUPER-RESOLUTION | 1700 |
| <i>Xinrui Jiang, Nannan Wang, Jingwei Xin, Keyu Li, Xi Yang, Xinbo Gao</i> | |
| WHAT TO SELECT: PURSUING CONSISTENT MOTION SEGMENTATION FROM MULTIPLE GEOMETRIC MODELS | 1708 |
| <i>Yangbangyan Jiang, Qianqian Xu, Ke Ma, Zhiyong Yang, Xiaochun Cao, Qingming Huang</i> | |
| ASYNCHRONOUS TEACHER GUIDED BIT-WISE HARD MINING FOR ONLINE HASHING | 1717 |
| <i>Sheng Jin, Qin Zhou, Hongxun Yao, Yao Liu, Xian-Sheng Hua</i> | |
| DEEP LOW-CONTRAST IMAGE ENHANCEMENT USING STRUCTURE TENSOR REPRESENTATION | 1725 |
| <i>Hyungjoo Jung, Hyunsung Jang, Namkoo Ha, Kwanghoon Sohn</i> | |
| SPECTRAL DISTRIBUTION AWARE IMAGE GENERATION | 1734 |
| <i>Steffen Jung, Margret Keuper</i> | |
| STARNET: TOWARDS WEAKLY SUPERVISED FEW-SHOT OBJECT DETECTION | 1743 |
| <i>Leonid Karlinsky, Joseph Shtok, Amit Alfassy, Moshe Lichtenstein, Sivan Harary, Eli Schwartz, Sivan Doveh, Prasanna Sattigeri, Rogerio Feris, Alex Bronstein, Raja Giryes</i> | |
| DISCRIMINATIVE REGION SUPPRESSION FOR WEAKLY-SUPERVISED SEMANTIC SEGMENTATION | 1754 |
| <i>Beomyoung Kim, Sangeun Han, Junmo Kim</i> | |
| VISUAL COMFORT AWARE-REINFORCEMENT LEARNING FOR DEPTH ADJUSTMENT OF STEREOSCOPIC 3D IMAGES | 1762 |
| <i>Hak Gu Kim, Minh Park, Sangmin Lee, Seongyeop Kim, Yong Man Ro</i> | |
| DUAL COMPOSITIONAL LEARNING IN INTERACTIVE IMAGE RETRIEVAL | 1771 |
| <i>Jongseok Kim, Youngjae Yu, Hoeseong Kim, Gunhee Kim</i> | |

| | |
|---|------|
| END-TO-END DIFFERENTIABLE LEARNING TO HDR IMAGE SYNTHESIS FOR MULTI-EXPOSURE IMAGES | 1780 |
| <i>Junghee Kim, Siyeong Lee, Suk-Ju Kang</i> | |

| | |
|---|------|
| STRUCTURED CO-REFERENCE GRAPH ATTENTION FOR VIDEO-GROUNDED DIALOGUE | 1789 |
| <i>Junyeong Kim, Sunjae Yoon, Dahyun Kim, Chang D. Yoo</i> | |

Author Index