

15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2020)

Volume 5: VISAPP

Valletta, Malta
27 – 29 February 2020

Editors:

**Giovanni Maria Farinella
Petia Radeva
Jose Braz**

ISBN: 978-1-7138-4044-2

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© (2020) by SCITEPRESS – Science and Technology Publications, Lda.
All rights reserved.

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

For permission requests, please contact SCITEPRESS – Science and Technology Publications, Lda.
at the address below.

SCITEPRESS – Science and Technology Publications, Lda.
Avenida de S. Francisco Xavier, Lote 7 Cv. C,
2900-616 Setúbal, Portugal

Phone: +351 265 520 185

Fax: +351 265520 186

info@scitepress.org

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

CONTENTS

INVITED SPEAKERS

KEYNOTE SPEAKERS

| | |
|--|----|
| Semantic 3D Scene Understanding in RGB-D Scans <i>Matthias Niessner</i> | 5 |
| The Computing Challenges of Mixed-Reality <i>Anthony Steed</i> | 7 |
| High Dynamic Range: Where to next? <i>Alan Chalmers</i> | 9 |
| Experimental Pitfalls <i>Helen Purchase</i> | 11 |

IMAGE AND VIDEO UNDERSTANDING

FULL PAPERS

| | |
|--|----|
| Using Local Refinements on 360 Stitching from Dual-fisheye Cameras <i>Rafael Roberto, Daniel Perazzo, João Paulo Lima, Veronica Teichrieb, Jonysberg Peixoto Quintino, Fabio Q. B. da Silva, Andre L. M. Santos and Helder Pinho</i> | 17 |
| Audio-guided Video Interpolation via Human Pose Features <i>Takayuki Nakatsuka, Masatoshi Hamanaka and Shigeo Morishima</i> | 27 |
| Semantic Scene Completion from a Single 360-Degree Image and Depth Map <i>Aloisio Dourado, Hansung Kim, Teofilo E. de Campos and Adrian Hilton</i> | 36 |
| FootAndBall: Integrated Player and Ball Detector <i>Jacek Komorowski, Grzegorz Kurzejamski and Grzegorz Sarwas</i> | 47 |
| Fine-tuning Siamese Networks to Assess Sport Gestures Quality <i>Mégane Millan and Catherine Achard</i> | 57 |
| Pre- and Post-processing Strategies for Generic Slice-wise Segmentation of Tomographic 3D Datasets Utilizing U-Net Deep Learning Models Trained for Specific Diagnostic Domains <i>Gerald A. Zwettler, Werner Backfrieder and David R. Holmes III</i> | 66 |
| Estimation of Muscle Fascicle Orientation in Ultrasonic Images <i>Regina Pohle-Fröhlich, Christoph Dalitz, Charlotte Richter, Tobias Hahnen, Benjamin Stäudle and Kirsten Albracht</i> | 79 |
| Deep Learning for Astronomical Object Classification: A Case Study <i>Ana Martinazzo, Mateus Espadoto and Nina S. T. Hirata</i> | 87 |
| Semi-automatic Learning Framework Combining Object Detection and Background Subtraction <i>Sugino Nicolas Alejandro, Tsubasa Minematsu, Atsushi Shimada, Takashi Shibata, Rin-ichiro Taniguchi, Eiji Kaneko and Hiroyoshi Miyano</i> | 96 |

| | |
|---|-----|
| Anomaly Event Detection based on People Trajectories for Surveillance Videos <i>Rensso Mora Colque, Edward Cayllahua, Victor C. de Melo, Guillermo Camara Chavez and William Robson Schwartz</i> | 107 |
| Fuzzy Fusion for Two-stream Action Recognition <i>Anderson Carlos Sousa e Santos, Helena de Almeida Maia, Marcos Roberto e Souza, Marcelo Bernardes Vieira and Helio Pedrini</i> | 117 |
| Coarse to Fine Vertebrae Localization and Segmentation with Spatial Configuration-Net and U-Net <i>Christian Payer, Darko Štern, Horst Bischof and Martin Urschler</i> | 124 |
| Self-supervised Depth Estimation based on Feature Sharing and Consistency Constraints <i>Julio Mendoza and Helio Pedrini</i> | 134 |
| RaDE: A Rank-based Graph Embedding Approach <i>Filipe Alves de Fernando, Daniel Carlos Guimarães Pedronette, Gustavo José de Sousa, Lucas Pascotti Valem and Ivan Rizzo Guilherme</i> | 142 |
| 3D Plant Growth Prediction via Image-to-Image Translation <i>Tomohiro Hamamoto, Hideaki Uchiyama, Atsushi Shimada and Rin-ichiro Taniguchi</i> | 153 |
| ConvPoseCNN: Dense Convolutional 6D Object Pose Estimation <i>Catherine Capellen, Max Schwarz and Sven Behnke</i> | 162 |
| Quantitative Analysis of Facial Paralysis using GMM and Dynamic Kernels <i>Nazil Perveen, Chalavadi Krishna Mohan and Yen Wei Chen</i> | 173 |
| Improving Dense Crowd Counting Convolutional Neural Networks using Inverse k-Nearest Neighbor Maps and Multiscale Upsampling <i>Greg Olmschenk, Hao Tang and Zhigang Zhu</i> | 185 |
| Geometric Deep Learning on Skeleton Sequences for 2D/3D Action Recognition <i>Rasha Friji, Hassen Drira and Faten Chaieb</i> | 196 |
| Mini V-Net: Depth Estimation from Single Indoor-Outdoor Images using Strided-CNN <i>Ahmed J. Afifi, Olaf Hellwich and Toufique Ahmed Soomro</i> | 205 |
| SHORT PAPERS | |
| A CNN-based Feature Space for Semi-supervised Incremental Learning in Assisted Living Applications <i>Tobias Scheck, Ana Perez Grassi and Gangolf Hirtz</i> | 217 |
| Towards Deep People Detection using CNNs Trained on Synthetic Images <i>Roberto Martín-López, David Fuentes-Jiménez, Sara Luengo-Sánchez, Cristina Losada-Gutiérrez, Marta Marrón-Romera and Carlos Luna</i> | 225 |
| Resources and End-to-End Neural Network Models for Arabic Image Captioning <i>Obeida ElJundi, Mohamad Dhaybi, Kotaiba Mokadam, Hazem Hajj and Daniel Asmar</i> | 233 |
| Multi-Branch Convolutional Descriptors for Content-based Remote Sensing Image Retrieval <i>Raffaele Imbriaco, Tunc Alkanat, Egor Bondarev and Peter H. N. de With</i> | 242 |
| SSD-ML: Hierarchical Object Classification for Traffic Surveillance <i>M. H. Zwemer, R. G. J. Wijnhoven and P. H. N. de With</i> | 250 |

| | |
|---|-----|
| Robust Method for Detecting Defect in Images Printed on 3D Micro-textured Surfaces: Modified Multiple Paired Pixel Consistency <i>Sheng Xiang, Shun'ichi Kaneko and Dong Liang</i> | 260 |
| Vessel-speed Enforcement System by Multi-camera Detection and Re-identification <i>H. G. J. Groot, M. H. Zwemer, R. G. J. Wijnhoven, Y. Bondarev and P. H. N. de With</i> | 268 |
| Semisupervised Classification of Anomalies Signatures in Electrical Wafer Sorting (EWS) Maps <i>Luigi C. Viagrande, Filippo L. M. Milotta, Paola Giuffrè, Giuseppe Bruno, Daniele Vinciguerra and Giovanni Gallo</i> | 278 |
| A Deep Transfer Learning Framework for Pneumonia Detection from Chest X-ray Images <i>Kh Tohidul Islam, Sudanthi Wijewickrema, Aaron Collins and Stephen O'Leary</i> | 286 |
| Facial Expression Recognition using the Bilinear Pooling <i>Marwa Ben Jabra, Ramzi Guetari, Aladine Chetouani, Hedi Tabia and Nawres Khelifa</i> | 294 |
| Acquisition of Optimal Connection Patterns for Skeleton-based Action Recognition with Graph Convolutional Networks <i>Katsutoshi Shiraki, Tsubasa Hirakawa, Takayoshi Yamashita and Hironobu Fujiyoshi</i> | 302 |
| Clothing Category Classification using Common Models Adaptively Adjusted to Observation <i>Jingyu Hu, Nobuyuki Kita and Yasuyo Kita</i> | 310 |
| Deep Body-pose Estimation via Synthetic Depth Data: A Case Study <i>Christopher Pramerdorfer and Martin Kampel</i> | 318 |
| Detecting Anomalous Regions from an Image based on Deep Captioning <i>Yusuke Hatae, Qingpu Yang, Muhammad Fikko Fadmiriratno, Yuanyuan Li, Tetsu Matsukawa and Einoshin Suzuki</i> | 326 |
| FSSSD: Fixed Scale SSD for Vehicle Detection <i>Jiwon Jun, Hyunjeong Pak and Moongu Jeon</i> | 336 |
| MobText: A Compact Method for Scene Text Localization <i>Luis Gustavo Lorgus Decker, Allan da Silva Pinto, Jose Luis Flores Campana, Manuel Cordova Neira, Andreza A. dos Santos, Jhonatas S. Conceição, Marcus A. Angeloni, Lin Tzy Li and Ricardo da S. Torres</i> | 343 |
| Multi-stream Architecture with Symmetric Extended Visual Rhythms for Deep Learning Human Action Recognition <i>Hemerson Tacon, André de Souza Brito, Hugo de Lima Chaves, Marcelo Bernardes Vieira, Saulo Moraes Villela, Helena de Almeida Maia, Darwin Tritto Concha and Helio Pedrini</i> | 351 |
| Monocular 3D Head Reconstruction via Prediction and Integration of Normal Vector Field <i>Oussema Bouafff, Bogdan Khomutenko and Mohamed Daoudi</i> | 359 |
| Time-unfolding Object Existence Detection in Low-quality Underwater Videos using Convolutional Neural Networks <i>Helmut Tödtmann, Matthias Vahl, Uwe Freiherr von Lukas and Torsten Ullrich</i> | 370 |
| Explaining Spatial Relation Detection using Layerwise Relevance Propagation <i>Gabriel Farrugia and Adrian Muscat</i> | 378 |
| Encrypted Image Display based on Individual Visual Characteristics <i>Ryota Niwa, Fumihiko Sakaue and Jun Sato</i> | 386 |

| | |
|--|-----|
| Detecting and Locating Boats using a PTZ Camera with Both Optical and Thermal Sensors <i>Christoffer P. Simonsen, Frederik M. Thiesson, Øyvind Holtsgo and Rikke Gade</i> | 395 |
| Manifold Learning-based Clustering Approach Applied to Anomaly Detection in Surveillance Videos <i>Leonardo Tadeu Lopes, Lucas Pascotti Valem, Daniel Carlos Guimarães Pedronette, Ivan Rizzo Guilherme, João Paulo Papa, Marcos Cleison Silva Santana and Danilo Colombo</i> | 404 |
| Multi-stream Deep Networks for Vehicle Make and Model Recognition <i>Mohamed Dhia Elhak Besbes, Yousri Kessentini and Hedi Tabia</i> | 413 |
| Texture-based 3D Face Recognition using Deep Neural Networks for Unconstrained Human-machine Interaction <i>Michael Danner, Patrik Huber, Muhammad Awais, Zhen-Hua Feng, Josef Kittler and Matthias Raetsch</i> | 420 |
| CNN Hyperparameter Tuning Applied to Iris Liveness Detection <i>Gabriela Yukari Kimura, Diego Rafael Lucio, Alceu S. Britto Jr. and David Menotti</i> | 428 |
| SOANets: Encoder-decoder based Skeleton Orientation Alignment Network for White Cane User Recognition from 2D Human Skeleton Sequence <i>Naoki Nishida, Yasutomo Kawanishi, Daisuke Deguchi, Ichiro Ide, Hiroshi Murase and Jun Piao</i> | 435 |
| Visual Descriptor Learning from Monocular Video <i>Umashankar Deekshith, Nishit Gajjar, Max Schwarz and Sven Behnke</i> | 444 |
| A New Approach Combining Trained Single-view Networks with Multi-view Constraints for Robust Multi-view Object Detection and Labelling <i>Yue Zhang, Adrian Hilton and Jean-Yves Guillemaut</i> | 452 |
| A Method for Detecting Human-object Interaction based on Motion Distribution around Hand <i>Tatsuhiko Tsukamoto, Toru Abe and Takuo Suganuma</i> | 462 |
| A Multi-purpose RGB-D Dataset for Understanding Everyday Objects <i>Shuichi Akizuki and Manabu Hashimoto</i> | 470 |
| Dynamic Mode Decomposition via Dictionary Learning for Foreground Modeling in Videos <i>Israr Ul Haq, Keisuke Fujii and Yoshinobu Kawahara</i> | 476 |
| Anomaly Detection in Surveillance Videos by Future Appearance-motion Prediction <i>Tuan-Hung Vu, Sebastien Ambellouis, Jacques Boonaert and Abdelmalik Taleb-Ahmed</i> | 484 |
| Distributed Information Integration in Convolutional Neural Networks <i>Dinesh Kumar and Dharmendra Sharma</i> | 491 |
| Plant Species Identification using Discriminant Bag of Words (DBoW) <i>Fiza Murtaza, Umber Saba, Muhammad Haroon Yousaf and Serestina Viriri</i> | 499 |
| Classification of Histopathological Images using Scale-Invariant Feature Transform <i>Andrzej Bukala, Bogusław Cyganek, Michał Koziarski, Bogdan Kwolek, Bogusław Olborski, Zbigniew Antosz, Jakub Swadźba and Piotr Sitkowski</i> | 506 |
| Proxy Embeddings for Face Identification among Multi-Pose Templates <i>Weronika Gutfeter and Andrzej Pacut</i> | 513 |
| Who Loves Virtue as much as He Loves Beauty?: Deep Learning based Estimator for Aesthetics of Portraits <i>Tobias Gerlach, Michael Danner, Le Ping Peng, Aidas Kaminickas, Wu Fei and Matthias Rättsch</i> | 521 |

| | |
|---|-----|
| DA-NET: Monocular Depth Estimation using Disparity Maps Awareness NETWORK <i>Antoine Billy and Pascal Desbarats</i> | 529 |
| Mask-guided Image Classification with Siamese Networks <i>Hiba Alqasir, Damien Muselet and Christophe Ducottet</i> | 536 |
| Light Field Scattering in Participating Media <i>Takuya Mokutani, Fumihiko Sakaue and Jun Sato</i> | 544 |
| Automated Generation of Synthetic in-Car Dataset for Human Body Pose Detection <i>João Borges, Bruno Oliveira, Helena Torres, Nelson Rodrigues, Sandro Queirós, Maximilian Shiller, Victor Coelho, Johannes Pallauf, José Henrique Brito, José Mendes and Jaime C. Fonseca</i> | 550 |
| Multimodal Dance Recognition <i>Monika Wysoczkańska and Tomasz Trzciński</i> | 558 |
| Towards the Automatic Visual Monitoring of Electricity Pylons from Aerial Images <i>Anicetus Odo, Stephen McKenna, David Flynn and Jan Vorstius</i> | 566 |
| Towards Unsupervised Image Harmonisation <i>Alan Dolhasz, Carlo Harvey and Ian Williams</i> | 574 |
| A Preliminary Study on the Automatic Visual based Identification of UAV Pilots from Counter UAVs <i>Dario Cazzato, Claudio Cimarrelli and Holger Voos</i> | 582 |
| 3D Convolutional Neural Network for Falling Detection using Only Depth Information <i>Sara Luengo Sánchez, Sergio de López Diz, David Fuentes-Jiménez, Cristina Losada-Gutiérrez, Marta Marrón-Romera and Ibrahim Sarker</i> | 590 |
| Hidden Markov Models for Pose Estimation <i>László Czúni and Amr M. Nagy</i> | 598 |
| MOBILE AND EGOCENTRIC VISION FOR HUMANS AND ROBOTS | |
| FULL PAPER | |
| Localizing Visitors in Natural Sites Exploiting Modality Attention on Egocentric Images and GPS Data <i>Giovanni Pasqualino, Stefano Scafiti, Antonino Furnari and Giovanni Maria Farinella</i> | 609 |
| SHORT PAPERS | |
| Geo-localization using Ridgeline Features Extracted from 360-degree Images of Sand Dunes <i>Shogo Fukuda, Shintaro Nakatani, Masashi Nishiyama and Yoshio Iwai</i> | 621 |
| A Comparison of Visual Navigation Approaches based on Localization and Reinforcement Learning in Virtual and Real Environments <i>Marco Rosano, Antonino Furnari, Luigi Gulino and Giovanni Maria Farinella</i> | 628 |
| Regression-based 3D Hand Pose Estimation using Heatmaps <i>Chaitanya Bandi and Ulrike Thomas</i> | 636 |
| CAD-based Learning for Egocentric Object Detection in Industrial Context <i>Julia Cohen, Carlos Crispim-Junior, Céline Grange-Faivre and Laure Tougne</i> | 644 |

| | |
|---|-----|
| Monocular 3D Object Detection via Geometric Reasoning on Keypoints <i>Ivan Barabanau, Alexey Artemov, Evgeny Burnaev and Vyacheslav Murashkin</i> | 652 |
| A First-person Database for Detecting Barriers for Pedestrians <i>Zenonas Theodosiou, Harris Partaourides, Tolga Atun, Simoni Panayi and Andreas Lanitis</i> | 660 |
| Towards Visual Loop Detection in Underwater Robotics using a Deep Neural Network <i>Antoni Burguera and Francisco Bonin-Font</i> | 667 |
| Dynamic Detectors of Oriented Spatial Contrast from Isotropic Fixational Eye Movements <i>Simone Testa, Giacomo Indiveri and Silvio P. Sabatini</i> | 674 |
| Evaluation of 3D Vision Systems for Detection of Small Objects in Agricultural Environments <i>Justin Le Louedec, Bo Li and Grzegorz Cielniak</i> | 682 |
| 3D Model-based 6D Object Pose Tracking on RGB Images using Particle Filtering and Heuristic Optimization <i>Mateusz Majcher and Bogdan Kwolek</i> | 690 |
| APPLICATIONS AND SERVICES | |
| FULL PAPERS | |
| Melanoma Detection System based on a Game Theory Model <i>Djamila Dahmani, Slimane Larabi, Sihame Djelouah, Nafissa Benhebbadj and Mehdi Cheref</i> | 703 |
| Towards Detecting Simultaneous Fear Emotion and Deception Behavior in Speech <i>Safa Chebbi and Sofia Ben Jebara</i> | 712 |
| Improving Age Estimation in Minors and Young Adults with Occluded Faces to Fight Against Child Sexual Exploitation <i>Deisy Chaves, Eduardo Fidalgo, Enrique Alegre, Francisco Jáñez-Martino and Rubel Biswas</i> | 721 |
| Multi-view Data Capture using Edge-synchronised Mobiles <i>Matteo Bortolon, Paul Chippendale, Stefano Messelodi and Fabio Poiesi</i> | 730 |
| SHORT PAPERS | |
| CHAKEL-DB: Online Database for Handwriting Diacritic Arabic Character <i>Houda Nakkach, Sofiene Haboubi and Hamid Amiri</i> | 743 |
| Human Climbing and Bouldering Motion Analysis: A Survey on Sensors, Motion Capture, Analysis Algorithms, Recent Advances and Applications <i>Julia Richter, Raul Beltrán Beltrán, Guido Köstermeyer and Ulrich Heinkel</i> | 751 |
| Detection, Counting and Maturity Assessment of Cherry Tomatoes using Multi-spectral Images and Machine Learning Techniques <i>I-Tzu Chen and Huei-Yung Lin</i> | 759 |
| Stereoscopic Text-based CAPTCHA on Head-Mounted Displays <i>Tadaaki Hosaka and Shinnosuke Furuya</i> | 767 |
| Vehicle Detection and Classification in Aerial Images using Convolutional Neural Networks <i>Chih-Yi Li and Huei-Yung Lin</i> | 775 |

| | |
|--|-----|
| PlanAR: Accurate and Stable 3D Positioning System via Interactive Plane Reconstruction for Handheld Augmented Reality <i>Ami Miyake, Hideaki Uchiyama, Atsushi Shimada and Rin-ichiro Taniguchi</i> | 783 |
| Acquisition Evaluation on Outdoor Scanning for Archaeological Artifact Digitalization <i>Aufaclav Zatu Kusuma Frisky, Adieyatna Fajri, Simon Brenner and Robert Sablatnig</i> | 792 |
| MAUI: Tele-assistance for Maintenance of Cyber-physical Systems <i>Philipp Fleck, Fernando Reyes-Aviles, Christian Pirchheim, Clemens Arth and Dieter Schmalstieg</i> | 800 |
| Player Tracking using Multi-viewpoint Images in Basketball Analysis <i>Shuji Tanikawa and Norio Tagawa</i> | 813 |
| Steganalysis of Semi-fragile Watermarking Systems Resistant to JPEG Compression <i>Anna Egorova and Victor Fedoseev</i> | 821 |
| Pectoral Muscle Segmentation in Tomosynthesis Images using Geometry Information and Grey Wolf Optimizer <i>Mohamed Abdel-Nasser, Francesc Porta Solsona and Domenec Puig</i> | 829 |
| Surgery Recording without Occlusions by Multi-view Surgical Videos <i>Tomohiro Shimizu, Kei Oishi, Ryo Hachiuma, Hiroki Kajita, Yoshihumi Takatsume and Hideo Saito</i> | 837 |
| Automatic Classification of Cervical Cell Patches based on Non-geometric Characteristics <i>Douglas Wender A. Isidoro, Cláudia M. Carneiro, Mariana T. Resende, Fátima N. S. Medeiros, Daniela M. Ushizima and Andrea G. Campos Bianchi</i> | 845 |
| Fully Connected Visual Words for the Classification of Skin Cancer Confocal Images <i>Athanasios Kallipolitis, Alexandros Stratigos, Alexios Zarras and Ilias Maglogiannis</i> | 853 |
| Driving Video Prediction based on Motion Estimation of 3D Objects using a Stereo Camera System <i>Takuya Umemura, Fumihiko Sakaue and Jun Sato</i> | 859 |
| AUTHOR INDEX | 867 |