

# **17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022)**

Volume 5: VISAPP

Online  
6 - 8 February 2022

**Editors:**

**Giovanni Maria Farinella  
Petia Radeva  
Kadi Bouatouch**

ISBN: 978-1-7138-5315-2

**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© (2022) by SCITEPRESS – Science and Technology Publications, Lda.  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2022)

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](mailto: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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# CONTENTS

---

## INVITED SPEAKERS

### KEYNOTE SPEAKERS

Wearable Human Augmentation <i>Roope Raisamo</i>	5
Reducing Brain-computer Interaction Training Time with Embodied Virtual Avatar <i>Filip Škola and Fotis Liarokapis</i>	7
The Risky Business of Visualizing Known Unknowns for Decision Making with Maps <i>Sara Irina Fabrikant</i>	19
Neural Implicit Representations for 3D Vision and Beyond <i>Andreas Geiger</i>	21

### IMAGE AND VIDEO UNDERSTANDING

#### FULL PAPERS

Self-supervised Learning from Semantically Imprecise Data <i>Clemens-Alexander Brust, Björn Barz and Joachim Denzler</i>	27
Animal Fiber Identification under the Open Set Condition <i>Oliver Rippel, Sergen Gülçelik, Khosrow Rahimi, Juliana Kurniadi, Andreas Herrmann and Dorit Merhof</i>	36
Who Did It? Identifying Foul Subjects and Objects in Broadcast Soccer Videos <i>Chunbo Song and Christopher Rasmussen</i>	48
CGT: Consistency Guided Training in Semi-Supervised Learning <i>Nesreen Hasan, Farzin Ghorban, Jörg Velten and Anton Kummert</i>	55
Unidentified Floating Object Detection in Maritime Environment <i>Darshan Venkatrayappa, Agnès Desolneux, Jean-Michel Hubert and Josselin Manceau</i>	65
Milking CowMask for Semi-supervised Image Classification <i>Geoff French, Avital Oliver and Tim Salimans</i>	75
Anomaly Detection for Industrial Inspection using Convolutional Autoencoder and Deep Feature-based One-class Classification <i>Jamal Saeedi and Alessandro Giusti</i>	85
3D Detection of Vehicles from 2D Images in Traffic Surveillance <i>M. H. Zwemer, D. Scholte, R. G. J. Wijnhoven and P. H. N. de With</i>	97
Iterative 3D Deformable Registration from Single-view RGB Images using Differentiable Rendering <i>Arul Selvam Periyasamy, Max Schwarz and Sven Behnke</i>	107
SieveNet: Estimating the Particle Size Distribution of Kernel Fragments in Whole Plant Corn Silage <i>Christoffer Bøgelund Rasmussen, Kristian Kirk and Thomas B. Moeslund</i>	117

GAN-based Face Mask Removal using Facial Landmarks and Pixel Errors in Masked Region <i>Hitoshi Yoshihashi, Naoto Ienaga and Maki Sugimoto</i>	125
Fine-grained Action Recognition using Attribute Vectors <i>Sravani Yenduri, Nazil Perveen, Vishnu Chalavadi and C. Krishna Mohan</i>	134
Hear Me out: Fusional Approaches for Audio Augmented Temporal Action Localization <i>Anurag Bagchi, Jazib Mahmood, Dolton Fernandes and Ravi Kiran Sarvadevabhatla</i>	144
SemSegDepth: A Combined Model for Semantic Segmentation and Depth Completion <i>Juan Pablo Lagos and Esa Rahtu</i>	155
On the Cross-dataset Generalization in License Plate Recognition <i>Rayson Laroca, Everton V. Cardoso, Diego R. Lucio, Valter Estevam and David Menotti</i>	166
Leveraging Local Domains for Image-to-Image Translation <i>Anthony Dell'Eva, Fabio Pizzati, Massimo Bertozzi and Raoul de Charette</i>	179
CAM-SegNet: A Context-Aware Dense Material Segmentation Network for Sparsely Labelled Datasets <i>Yuwen Heng, Yihong Wu, Srinandan Dasmahapatra and Hansung Kim</i>	190
The MVTec 3D-AD Dataset for Unsupervised 3D Anomaly Detection and Localization <i>Paul Bergmann, Xin Jin, David Sattlegger and Carsten Steger</i>	202
Improving the Sample-complexity of Deep Classification Networks with Invariant Integration <i>Matthias Rath and Alexandru Paul Condurache</i>	214
Segregational Soft Dynamic Time Warping and Its Application to Action Prediction <i>Victoria Manousaki and Antonis Argyros</i>	226
Vehicle Pair Activity Classification using QTC and Long Short Term Memory Neural Network <i>Rahuln Radhakrishnan and Alaa Alzoubi</i>	236
ETL: Efficient Transfer Learning for Face Tasks <i>Thrupthi Ann John, Isha Dua, Vineeth N. Balasubramanian and C. V. Jawahar</i>	248
Reinforced Damage Minimization in Critical Events for Self-driving Vehicles <i>Francesco Merola, Fabrizio Falchi, Claudio Gennaro and Marco Di Benedetto</i>	258
Evolutional Normal Maps: 3D Face Representations for 2D-3D Face Recognition, Face Modelling and Data Augmentation <i>Michael Danner, Thomas Weber, Patrik Huber, Muhammad Awais, Matthias Raetsch and Josef Kittler</i>	267
Bispectral Pedestrian Detection Augmented with Saliency Maps using Transformer <i>Mohamed Amine Marnissi, Ikram Hattab, Hajer Fradi, Anis Sahbani and Najoua Essoukri Ben Amara</i>	275

## SHORT PAPERS

Weakly-supervised Localization of Multiple Objects in Images using Cosine Loss <i>Björn Barz and Joachim Denzler</i>	287
DeepPupil Net: Deep Residual Network for Precise Pupil Center Localization <i>Nikolaos Pouloupoulos and Emmanouil Z. Psarakis</i>	297
High Resolution Mask R-CNN-based Damage Detection on Titanium Nitride Coated Milling Tools for Condition Monitoring by using a New Illumination Technique <i>Mühenad Bilal, Sunil Kancharana, Christian Mayer, Daniel Pfaller, Leonid Koval, Markus Bregulla, Rafal Cupek and Adam Ziębiński</i>	305
Neural Network Pruning based on Filter Importance Values Approximated with Monte Carlo Gradient Estimation <i>Csanád Sándor, Szabolcs Pável and Lehel Csató</i>	315
The MIS Check-Dam Dataset for Object Detection and Instance Segmentation Tasks <i>Chintan Tundia, Rajiv Kumar, Om Damani and G. Sivakumar</i>	323
Oil Spill Detection and Visualization from UAV Images using Convolutional Neural Networks <i>Valério N. Rodrigues Junior, Roberto J. M. Cavalcante, João A. G. R. Almeida, Tiago P. M. Fé, Ana C. M. Malhado, Thales Vieira and Krerley Oliveira</i>	331
Generating Proposals from Corners in RPN to Detect Bees in Dense Scenes <i>Yassine Kriouile, Corinne Ancourt, Katarzyna Wegrzyn-Wolska and Lamine Bougueroua</i>	339
Feature Extraction using Downsampling for Person Re-identification with Low-resolution Images <i>Masashi Nishiyama, Takuya Endo and Yoshio Iwai</i>	351
Applying Center Loss to Multidimensional Feature Space in Deep Neural Networks for Open-set Recognition <i>Daiju Kanaoka, Yuichiro Tanaka and Hakaru Tamukoh</i>	359
Automated Human Movement Segmentation by Means of Human Pose Estimation in RGB-D Videos for Climbing Motion Analysis <i>Raul Beltrán B., Julia Richter and Ulrich Heinkel</i>	366
A General Two-branch Decoder Architecture for Improving Encoder-decoder Image Segmentation Models <i>Sijie Hu, Fabien Bonardi, Samia Bouchafa and Désiré Sidibé</i>	374
Automated Damage Inspection of Power Transmission Towers from UAV Images <i>Aleixo Cambeiro Barreiro, Clemens Seibold, Anna Hilsmann and Peter Eisert</i>	382
Monocular Estimation of Translation, Pose and 3D Shape on Detected Objects using a Convolutional Autoencoder <i>Ivar Persson, Martin Ahrnbom and Mikael Nilsson</i>	390
Tennis Strokes Recognition from Generated Stick Figure Video Overlays <i>Boris Bačić and Ishara Bandara</i>	397
PodNet: Ensemble-based Classification of Podocytopathy on Kidney Glomerular Images <i>George Oliveira Barros, David Campos Wanderley, Luciano Oliveira Rebouças, Washington L. C. dos Santos, Angelo A. Duarte and Flavio de Barros Vidal</i>	405

Detecting Patches on Road Pavement Images Acquired with 3D Laser Sensors using Object Detection and Deep Learning <i>Syed Ibrahim Hassan, Dympna O'sullivan, Susan Mckeever, David Power, Ray Mcgowan and Kieran Feighan</i>	413
Weakly Supervised Segmentation of Histopathology Images: An Insight in Feature Maps Ability for Learning Models Interpretation <i>Yanbo Feng, Adel Hafiane and H�el�ene Laurent</i>	421
Multitask Metamodel for Keypoint Visibility Prediction in Human Pose Estimation <i>Romain Guesdon, Carlos Crispim-Junior and Laure Tougne</i>	428
Evaluation of Long-term Deep Visual Place Recognition <i>Farid Alijani, Jukka Peltom�aki, Jussi Puura, Heikki Huttunen, Joni-Kristian K�am�ar�ainen and Esa Rahtu</i>	437
Human Activity Recognition: A Spatio-temporal Image Encoding of 3D Skeleton Data for Online Action Detection <i>Nassim Mokhtari, Alexis N�ed�elec and Pierre De Loor</i>	448
Deep Set Conditioned Latent Representations for Action Recognition <i>Akash Singh, Tom de Schepper, Kevin Mets, Peter Hellinckx, Jos�e Oramas and Steven Latr�e</i>	456
Streamlining Action Recognition in Autonomous Shared Vehicles with an Audiovisual Cascade Strategy <i>Jo�o Ribeiro Pinto, Pedro Carvalho, Carolina Pinto, Afonso Sousa, Leonardo Capozzi and Jaime S. Cardoso</i>	467
Video-based Behavior Understanding of Children for Objective Diagnosis of Autism <i>Abid Ali, Farhood Negin, Susanne Th�ummler and Francois Bremond</i>	475
Buildings Extraction from Historical Topographic Maps via a Deep Convolution Neural Network <i>Christos Xydas, Anastasios L. Kesidis, Kleomenis Kalogeropoulos and Andreas Tsatsaris</i>	485
Subclass-based Undersampling for Class-imbalanced Image Classification <i>Daniel Lehmann and Marc Ebner</i>	493
Multimodal Personality Recognition using Cross-attention Transformer and Behaviour Encoding <i>Tanay Agrawal, Dhruv Agarwal, Michal Balazia, Neelabh Sinha and Fran�ois Bremond</i>	501
Improving Semantic Image Segmentation via Label Fusion in Semantically Textured Meshes <i>Florian Fervers, Timo Breuer, Gregor Stachowiak, Sebastian Bullinger, Christoph Bodensteiner and Michael Arens</i>	509
Honeybee Re-identification in Video: New Datasets and Impact of Self-supervision <i>Jeffrey Chan, Hector Carri�on, R�emi M�egret, Jos�e L. Agosto Rivera and Tugrul Giray</i>	517
Describing Image Focused in Cognitive and Visual Details for Visually Impaired People: An Approach to Generating Inclusive Paragraphs <i>Daniel L. Fernandes, Marcos H. F. Ribeiro, Fabio R. Cerqueira and Michel M. Silva</i>	526
Augmented Radar Points Connectivity based on Image Processing Techniques for Object Detection and Classification <i>Mohamed Sabry, Ahmed Hussein, Amr Elmougy and Slim Abdennadher</i>	535
Using Student Action Recognition to Enhance the Efficiency of Tele-education <i>Eleni Dimitriadou and Andreas Lanitis</i>	543

TVNet: Temporal Voting Network for Action Localization <i>Hanyuan Wang, Dima Damen, Majid Mirmehdi and Toby Perrett</i>	550
HRI-Gestures: Gesture Recognition for Human-Robot Interaction <i>Avgi Kollakidou, Frederik Haarslev, Cagatay Odabasi, Leon Bodenhagen and Norbert Krüger</i>	559
StructureNet: Deep Context Attention Learning for Structural Component Recognition <i>Akash Kaothalkar, Bappaditya Mandal and Niladri. B. Puhan</i>	567
Pose Guided Feature Learning for 3D Object Tracking on RGB Videos <i>Mateusz Majcher and Bogdan Kwolek</i>	574
Fusion of Different Features by Cross Cooperative Learning for Semantic Segmentation <i>Ryota Ikedo and Kazuhiro Hotta</i>	582
Detecting Object States vs Detecting Objects: A New Dataset and a Quantitative Experimental Study <i>Filippos Gouidis, Theodore Patkos, Antonis Argyros and Dimitris Plexousakis</i>	590
Skeleton-based Online Sign Language Recognition using Monotonic Attention <i>Natsuki Takayama, Gibran Benitez-Garcia and Hiroki Takahashi</i>	601
Seeing the Differences in Artistry among Art Fields by using Multi-task Learning <i>Ryo Sato, Fumihiko Sakaue and Jun Sato</i>	609
Diversifying Image Synthesis using Data Classification <i>Yuta Suzuki, Fumihiko Sakaue and Jun Sato</i>	617
Combining Text and Image Knowledge with GANs for Zero-Shot Action Recognition in Videos <i>Kaiqiang Huang, Luis Miralles-Pechuán and Susan Mckeever</i>	623
Exploitation of Noisy Automatic Data Annotation and Its Application to Hand Posture Classification <i>Georgios Lydakis, Iason Oikonomidis, Dimitrios Kosmopoulos and Antonis A. Argyros</i>	632
Transfer Learning via Test-time Neural Networks Aggregation <i>Bruno Casella, Alessio Barbaro Chisari, Sebastiano Battiato and Mario Valerio Giuffrida</i>	642
Evaluation of RGB and LiDAR Combination for Robust Place Recognition <i>Farid Alijani, Jukka Peltomäki, Jussi Puura, Heikki Huttunen, Joni-Kristian Kämäräinen and Esa Rahtu</i>	650
Using Contrastive Learning and Pseudolabels to Learn Representations for Retail Product Image Classification <i>Muktabh Mayank Srivastava</i>	659
Don't Miss the Fine Print! An Enhanced Framework to Extract Text from Low Resolution Images <i>Pranay Dugar, Aditya Vikram, Anirban Chatterjee, Kunal Banerjee and Vijay Agneeswaran</i>	664
Attention-based Gender Recognition on Masked Faces <i>Vincenzo Carletti, Antonio Greco, Alessia Saggese and Mario Vento</i>	672
Class-conditional Importance Weighting for Deep Learning with Noisy Labels <i>Bhalaji Nagarajan, Ricardo Marques, Marcos Mejia and Petia Radeva</i>	679

## MOTION, TRACKING AND STEREO VISION

### FULL PAPERS

- MOD SLAM: Mixed Method for a More Robust SLAM without Loop Closing 691  
*Thomas Belos, Pascal Monasse and Eva Dokladalova*
- Event-based Extraction of Navigation Features from Unsupervised Learning of Optic Flow Patterns 702  
*Paul Fricker, Tushar Chauhan, Christophe Hurter and Benoit R. Cottreau*
- Scan2Part: Fine-grained and Hierarchical Part-level Understanding of Real-World 3D Scans 711  
*Alexandr Notchenko, Vladislav Ishimtsev, Alexey Artemov, Vadim Selyutin, Emil Bogomolov and Evgeny Burnaev*
- Classification Performance of RanSaC Algorithms with Automatic Threshold Estimation 723  
*Clément Riu, Vincent Nozick, Pascal Monasse and Joachim Dehais*
- BRDF-based Irradiance Image Estimation to Remove Radiometric Differences for Stereo Matching 734  
*Kebin Peng, John Quarles and Kevin Desai*
- Combining Local and Global Pose Estimation for Precise Tracking of Similar Objects 745  
*Niklas Gard, Anna Hilsmann and Peter Eisert*
- Pushing the Efficiency of StereoNet: Exploiting Spatial Sparsity 757  
*Georgios Zampokas, Christos-Savvas Bouganis and Dimitrios Tzovaras*

### SHORT PAPERS

- 3-D Tennis Ball Trajectory Estimation from Multi-view Videos and Its Parameterization by Parabola Fitting 769  
*Kenta Sohara and Yasuyuki Sugaya*
- Seg2Pose: Pose Estimations from Instance Segmentation Masks in One or Multiple Views for Traffic Applications 777  
*Martin Ahrnbom, Ivar Persson and Mikael Nilsson*
- Motion-constrained Road User Tracking for Real-time Traffic Analysis 785  
*Nyan Bo Bo, Peter Veelaert and Wilfried Philips*
- 3D Object Reconstruction using Stationary RGB Camera 793  
*José G. dos S. Júnior, Gustavo C. R. Lima, Adam H. M. Pinto, João Paulo S. do M. Lima, Veronica Teichrieb, Jonysberg P. Quintino, Fabio Q. B. da Silva, Andre L. M. Santos and Helder Pinho*
- Error Evaluation of Semantic VSLAM Algorithms for Smart Farming 801  
*Adam Kalisz, Mingjun Sun, Jonas Gedschold, Tim Erich Wegner, Giovanni Del Galdo and Jörn Thielecke*
- Viewpoint-independent Single-view 3D Object Reconstruction using Reinforcement Learning 811  
*Seiya Ito, Byeongjun Ju, Naoshi Kaneko and Kazuhiko Sumi*
- Generalizable Online 3D Pedestrian Tracking with Multiple Cameras 820  
*Victor Lyra, Isabella de Andrade, João Paulo Lima, Rafael Roberto, Lucas Figueiredo, João Marcelo Teixeira, Diego Thomas, Hideaki Uchiyama and Veronica Teichrieb*
- LiMoSeg: Real-time Bird's Eye View based LiDAR Motion Segmentation 828  
*Sambit Mohapatra, Mona Hodaei, Senthil Yogamani, Stefan Milz, Heinrich Gotzig, Martin Simon, Hazem Rashed and Patrick Maeder*



Evaluating the Impact of Head Motion on Monocular Visual Odometry with Synthetic Data <i>Charles Hamesse, Hiep Luong and Rob Haelterman</i>	836
An Occlusion Aware Five-view Stereo System and Its Application in Video Post-production <i>Changan Zhu, Taryn Laurendeau and Chris Joslin</i>	844
Human Pose Estimation through a Novel Multi-view Scheme <i>Jorge L. Charco, Angel D. Sappa and Boris X. Vintimilla</i>	855
Visual Tilt Correction for Vehicle-mounted Cameras <i>Firas Kastantin and Levente Hajder</i>	863
Semi-supervised Anomaly Detection for Weakly-annotated Videos <i>Khaled El-Tahan and Marwan Torki</i>	871
AUTHOR INDEX	879