

19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2024)

Volume 2: VISAPP

Rome, Italy
27-29 February 2024

Editors:

**Petia Radeva
Antonino Furnari
Kadi Bouatouch
A. Augusto Sousa**

ISBN: 978-1-7138-9756-9

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© (2024) by SCITEPRESS – Science and Technology Publications, Lda.
All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

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

Towards Detailed Understanding of the Visual World <i>Fahad Khan</i>	5
Virtual Reality in Mental Health: A Self-Counselling Approach <i>Mel Slater</i>	7
Multi-Modal Human-Machine Interaction: Joint Optimization of Single Modalities and Automatic Learning of Communication Channel Fusion <i>Gerhard Rigoll</i>	9
The Dance of Logic and Unpredictability: Examining the Predictability of User Behavior on Visual Analytics Tasks <i>Alvitta Ottley</i>	11

IMAGE AND VIDEO UNDERSTANDING

FULL PAPERS

Enabling On-Device Continual Learning with Binary Neural Networks and Latent Replay <i>Lorenzo Vorabbi, Davide Maltoni, Guido Borghi and Stefano Santi</i>	25
Uncertainty-Based Detection of Adversarial Attacks in Semantic Segmentation <i>Kira Maag and Asja Fischer</i>	37
Synthesizing Classifiers from Prior Knowledge <i>G. J. Burghouts, K. Schutte, M. Kruithof, W. Huizinga, F. Ruis and H. Kuijff</i>	47
StyleHumanCLIP: Text-Guided Garment Manipulation for StyleGAN-Human <i>Takato Yoshikawa, Yuki Endo and Yoshihiro Kanamori</i>	59
S3Aug: Segmentation, Sampling, and Shift for Action Recognition <i>Taiki Sugiura and Toru Tamaki</i>	71
Attention-Based Shape and Gait Representations Learning for Video-Based Cloth-Changing Person Re-Identification <i>Vuong D. Nguyen, Samiha Mirza, Pranav Mantini and Shishir K. Shah</i>	80
Reducing Bias in Pre-Trained Models by Tuning While Penalizing Change <i>Niklas Penzel, Gideon Stein and Joachim Denzler</i>	90
Important Pixels Sampling for NeRF Training Based on Edge Values and Squared Errors Between the Ground Truth and the Estimated Colors <i>Kohei Fukuda, Takio Kurita and Hiroaki Aizawa</i>	102
Pixel-Wise Gradient Uncertainty for Convolutional Neural Networks Applied to Out-of-Distribution Segmentation <i>Kira Maag and Tobias Riedlinger</i>	112

Enabling RAW Image Classification Using Existing RGB Classifiers <i>Rasmus Munksø, Mathias Viborg Andersen, Lau Nørgaard, Andreas Møgelmoose and Thomas B. Moeslund</i>	123
Non-Local Context-Aware Attention for Object Detection in Remote Sensing Images <i>Yassin Terraf, El Mahdi Mercha and Mohammed Erradi</i>	130
Mediapi-RGB: Enabling Technological Breakthroughs in French Sign Language (LSF) Research Through an Extensive Video-Text Corpus <i>Yanis Ouakrim, Hannah Bull, Michèle Gouiffès, Denis Beautemps, Thomas Hueber and Annelies Braffort</i>	139
When Medical Imaging Met Self-Attention: A Love Story That Didn't Quite Work out <i>Tristan Piater, Niklas Penzel, Gideon Stein and Joachim Denzler</i>	149
RailCloud-HdF: A Large-Scale Point Cloud Dataset for Railway Scene Semantic Segmentation <i>Mahdi Abid, Mathis Teixeira, Ankur Mahtani and Thomas Laurent</i>	159
Investigating the Corruption Robustness of Image Classifiers with Random p -norm Corruptions <i>Georg Siedel, Weijia Shao, Silvia Vock and Andrey Morozov</i>	171
Calisthenics Skills Temporal Video Segmentation <i>Antonio Finocchiaro, Giovanni Maria Farinella and Antonino Furnari</i>	182
Detecting Anomalies in Textured Images Using Modified Transformer Masked Autoencoder <i>Afshin Dini and Esa Rahtu</i>	191
Parts-Based Implicit 3D Face Modeling <i>Yajie Gu and Nick Pears</i>	201
Robust Long-Tailed Image Classification via Adversarial Feature Re-Calibration <i>Jinghao Zhang, Zhenhua Feng and Yaochu Jin</i>	213
Alias-Free GAN for 3D-Aware Image Generation <i>Attila Szabó, Yevgeniy Puzikov, Sahan Ayyaz, Sonia Aurelio, Peter Gehler, Reza Shirvany and Malte Alf</i>	221
Conditional Vector Graphics Generation for Music Cover Images <i>Ivan Jarsky, Valeria Efimova, Ilya Bizyaev and Andrey Filchenkov</i>	233
Anomaly Detection on Roads Using an LSTM and Normal Maps <i>Yusuke Nonaka, Hideo Saito, Hideaki Uchiyama, Kyota Higa and Masahiro Yamaguchi</i>	244
Beyond the Known: Adversarial Autoencoders in Novelty Detection <i>Muhammad Asad, Ihsan Ullah, Ganesh Sistu and Michael G. Madden</i>	256
Image Generation from Hyper Scene Graphs with Trinomial Hyperedges Using Object Attention <i>Ryosuke Miyake, Tetsu Matsukawa and Einoshin Suzuki</i>	266
CSE: Surface Anomaly Detection with Contrastively Selected Embedding <i>Simon Thomine and Hichem Snoussi</i>	280

SHORT PAPERS

Improving Pseudo-Labeling and Enhancing Robustness for Semi-Supervised Domain Generalization <i>Adnan Khan, Mai A. Shaaban and Muhammad Haris Khan</i>	293
Synthetic Data-Driven Approach for Missing Nut and Bolt Classification in Flange Joints <i>Frankly Toro, Hassane Trigui, Yazeed Alnumay, Siddharth Mishra and Sahejad Patel</i>	301
Classification of Towels in a Robotic Workcell Using Deep Neural Networks <i>Jens Møller Rossen, Patrick Sjøgaard Terp, Norbert Krüger, Laus Skovgård Bigum and Tudor Morar</i>	309
Evaluating Learning Potential with Internal States in Deep Neural Networks <i>Shogo Takasaki and Shuichi Enokida</i>	317
Cybersecurity Intrusion Detection with Image Classification Model Using Hilbert Curve <i>Punyawat Jaroensiripong, Karin Sumongkayothin, Prarinya Siritanawan and Kazunori Kotani</i>	325
Pedestrian's Gaze Object Detection in Traffic Scene <i>Hiroto Murakami, Jialei Chen, Daisuke Deguchi, Takatsugu Hirayama, Yasutomo Kawanishi and Hiroshi Murase</i>	333
Lens Flare-Aware Detector in Autonomous Driving <i>Shanxing Ma and Jan Aelterman</i>	341
Improved Pest Detection in Insect Larvae Rearing with Pseudo-Labeling and Spatio-Temporal Masking <i>Paweł Majewski, Piotr Lampa, Robert Burduk and Jacek Reiner</i>	349
Online Human Activity Recognition Using Efficient Neural Architecture Search with Low Environmental Impact <i>Nassim Mokhtari, Alexis Nédélec, Marlène Gilles and Pierre De Loor</i>	357
Towards Rapid Prototyping and Comparability in Active Learning for Deep Object Detection <i>Tobias Riedlinger, Marius Schubert, Karsten Kahl, Hanno Gottschalk and Matthias Rottmann</i>	366
Deep Active Learning with Noisy Oracle in Object Detection <i>Marius Schubert, Tobias Riedlinger, Karsten Kahl and Matthias Rottmann</i>	375
Multi-Task Learning Based on Log Dynamic Loss Weighting for Sex Classification and Age Estimation on Panoramic Radiographs <i>Igor Prado, David Lima, Julian Liang, Ana Hougaz, Bernardo Peters and Luciano Oliveira</i>	385
A Comparative Evaluation of Self-Supervised Methods Applied to Rock Images Classification <i>Van Thao Nguyen, Dominique Fourer, Désiré Sidibé, Jean-François Lecomte and Souhail Youssef</i>	393
Sumo Action Classification Using Mawashi Keypoints <i>Yuki Utsuro, Hidehiko Shishido and Yoshinari Kameda</i>	401
Diffusion-Based Image Generation for In-Distribution Data Augmentation in Surface Defect Detection <i>Luigi Capogrosso, Federico Girella, Francesco Taioli, Michele Dalla Chiara, Muhammad Aqeel, Franco Fummi, Francesco Setti and Marco Cristani</i>	409
Class Weighted Focal Loss for Improving Class Imbalance in Semi-Supervised Object Detection <i>Shinichi Hoketsu, Tsubasa Hirakawa, Takayoshi Yamashita and Hironobu Fujiyoshi</i>	417

Kore Initial Clustering for Unsupervised Domain Adaptation <i>Kyungsik Lee, Youngmi Jun, Eunji Kim, Suhyun Kim, Seong Jae Hwang and Jonghyun Choi</i>	425
Interpretable Anomaly Analysis for Surveillance Video <i>Meng Dong</i>	433
Association of Grad-CAM, LIME and Multidimensional Fractal Techniques for the Classification of H&E Images <i>Thales R. S. Lopes, Guilherme F. Roberto, Carlos Soares, Thaína A. A. Tosta, Adriano B. Silva, Adriano M. Loyola, Sérgio V. Cardoso, Paulo R. de Faria, Marcelo Z. do Nascimento and Leandro A. Neves</i>	441
How Quality Affects Deep Neural Networks in Fine-Grained Image Classification <i>Joseph Smith, Zheming Zuo, Jonathan Stonehouse and Boguslaw Obara</i>	448
Comparative Study Between Object Detection Models, for Olive Fruit Fly Identification <i>Margarida Victoriano, Lino Oliveira and Hélder P. Oliveira</i>	458
Efficient Parameter Mining and Freezing for Continual Object Detection <i>Angelo G. Menezes, Augusto J. Peterlevitz, Mateus A. Chinelatto and André C. P.L. F. de Carvalho</i>	466
Body-Part Enabled Wildlife Detection and Tracking in Video Sequences <i>Alberto Lee, Kofi Appiah and Sze Chai Kwok</i>	475
Challenging the Black Box: A Comprehensive Evaluation of Attribution Maps of CNN Applications in Agriculture and Forestry <i>Lars Nieradzík, Henrike Stephani, Jördis Sieburg-Rockel, Stephanie Helmling, Andrea Olbrich and Janis Keuper</i>	483
GAF-Net: Video-Based Person Re-Identification via Appearance and Gait Recognitions <i>Moncef Boujou, Rabah Iguernaissi, Lionel Nicod, Djamel Merad and Séverine Dubuisson</i>	493
Applying Prompts and Parameter-Efficient Methods to Enhance Single-Stream Vision-Language Transformers <i>Xuehao Liu, Sarah Jane Delany and Susan McKeever</i>	501
Comprehensive Evaluation of End-to-End Driving Model Explanations for Autonomous Vehicles <i>Chenkai Zhang, Daisuke Deguchi, Jialei Chen and Hiroshi Murase</i>	509
Modification of DDIM Encoding for Generating Counterfactual Pathology Images of Malignant Lymphoma <i>Ryoichi Koga, Mauricio Kugler, Tatsuya Yokota, Kouichi Ohshima, Hiroaki Miyoshi, Miharuru Nagaishi, Noriaki Hashimoto, Ichiro Takeuchi and Hidekata Hontani</i>	519
Diverse Data Selection Considering Data Distribution for Unsupervised Continual Learning <i>Naoto Hayashi, Naoki Okamoto, Tsubasa Hirakawa, Takayoshi Yamashita and Hironobu Fujiyoshi</i>	528
Privacy-Preserving Face Recognition in Hybrid Frequency-Color Domain <i>Dong Han, Yong Li and Joachim Denzler</i>	536
Relationship Between Semantic Segmentation Model and Additional Features for 3D Point Clouds Obtained from on-Vehicle LIDAR <i>Hisato Hashimoto and Shuichi Enokida</i>	547
A Fusion Approach for Enhanced Remote Sensing Image Classification <i>Vian Abdulmajeed Ahmed, Khaled Jouini, Amel Tuama and Ouajdi Korbaa</i>	554

Agrinet: A Hyperspectral Image Based Precise Crop Classification Model <i>Aditi Palit, Himanshu Dolekar and Kalidas Yeturu</i>	562
Scale Learning in Scale-Equivariant Convolutional Networks <i>Mark Basting, Robert-Jan Bruintjes, Thaddäus Wiedemer, Matthias Kümmerer, Matthias Bethge and Jan van Gemert</i>	567
Unsupervised Few-Shot Image Segmentation with Dense Feature Learning and Sparse Clustering <i>Kuangdai Leng, Robert Atwood, Winfried Kockelmann, Deniza Chekrygina and Jeyan Thiyagalingam</i>	575
NeRF-Diffusion for 3D-Consistent Face Generation and Editing <i>Héctor Laria, Kai Wang, Joost van de Weijer, Bogdan Raducanu and Yaxing Wang</i>	587
Nearest Neighbor-Based Data Denoising for Deep Metric Learning <i>George Galanakis, Xenophon Zabulis and Antonis A. Argyros</i>	595
High Precision Single Shot Object Detection in Automotive Scenarios <i>Soumya A, C Krishna Mohan and Linga Reddy Cenkeramaddi</i>	604
Understanding Marker-Based Normalization for FLIM Networks <i>Leonardo de Melo Joao, Matheus Abrantes Cerqueira, Barbara Caroline Benato and Alexandre Xavier Falcao</i>	612
Ontology-Driven Deep Learning Model for Multitask Visual Food Analysis <i>Daniel Ponte, Eduardo Aguilar, Mireia Ribera and Petia Radeva</i>	624
AR-VPT: Simple Auto-Regressive Prompts for Adapting Frozen ViTs to Videos <i>Muhammad Zain Yousuf, Syed Talal Wasim, Syed Nouman Hasany and Muhammad Farhan</i>	632
Benchmarking Sensitivity of Continual Graph Learning for Skeleton-Based Action Recognition <i>Wei Wei, Tom De Schepper and Kevin Mets</i>	639
Multi-Task Planar Reconstruction with Feature Warping Guidance <i>Luan Wei, Anna Hilsmann and Peter Eisert</i>	652
A Cascade Methodology to Evaluate Black-Box Recognition Systems Based on a Copycat Algorithm <i>Dinh Cong Nguyen, Nhan Tam Le, Van Hoa Mai, Tuong Quan Nguyen, Van Quan Nguyen and The Cuong Nguyen</i>	662
Generating Videos from Stories Using Conditional GAN <i>Takahiro Kozaki, Fumihiko Sakaue and Jun Sato</i>	671
Anomaly Detection and Localization for Images of Running Paper Web in Paper Manufacturing <i>Afshin Dini, Marja Mettänen and Esa Rahtu</i>	678
Conic Linear Units: Improved Model Fusion and Rotational-Symmetric Generative Model <i>Changqing Fu and Laurent D. Cohen</i>	686
Unsupervised Annotation and Detection of Novel Objects Using Known Objectness <i>Harsh Singh Jadon, Jagdish Deshmukh, Kamakshya Prasad Nayak, Kamalakar Vijay Thakare and Debi Prasad Dogra</i>	694
Large Scale Graph Construction and Label Propagation <i>Z. Ibrahim, A. Bosaghzadeh and F. Dornaika</i>	702
Image Augmentation Preserving Object Parts Using Superpixels of Variable Granularity <i>D. Sun and F. Dornaika</i>	710

AbSynth: Using Abstract Image Synthesis for Synthetic Training <i>Dominik Penk, Maik Horn, Christoph Strohmeyer, Bernhard Egger, Marc Stamminger and Frank Bauer</i>	718
AV-PEA: Parameter-Efficient Adapter for Audio-Visual Multimodal Learning <i>Abduljalil Radman and Jorma Laaksonen</i>	730
Exploring the Impact of Knowledge Graphs on Zero-Shot Visual Object State Classification <i>Filippos Goudis, Konstantinos Papoutsakis, Theodore Patkos, Antonis Argyros and Dimitris Plexousakis</i>	738
Effects of Model Drift on Ship Detection Models <i>Namita Agarwal, Anh Vu Vo, Michela Bertolotto, Alan Barnett, Ahmed Khalid and Merry Globin</i>	750
Minimalist CNN for Medical Imaging Classification with Small Dataset: Does Size Really Matter and How? <i>Marie Économidès and Pascal Desbarats</i>	755
Occlusion-Robust and Efficient 6D Pose Estimation with Scene-Level Segmentation Refinement and 3D Partial-to-6D Full Point Cloud Transformation <i>Sukhan Lee, Soojin Lee and Yongjun Yang</i>	763
Identification of Honeybees with Paint Codes Using Convolutional Neural Networks <i>Gabriel Santiago-Plaza, Luke Meyers, Andrea Gomez-Jaime, Rafael Meléndez-Ríos, Fanfan Noel, Jose Agosto, Tugrul Giray, Josué Rodríguez-Cordero and Rémi Mégret</i>	772
Depth Estimation Using Weighted-Loss and Transfer Learning <i>Muhammad Adeel Hafeez, Michael G. Madden, Ganesh Sistu and Ihsan Ullah</i>	780
Strategies for Classifier Selection Based on Genetic Programming for Multimedia Data Recognition <i>Rafael Junqueira Martarelli, Douglas Rodrigues, Clayton Reginaldo Pereira, Jurandy Almeida and João Paulo Papa</i>	788
Enhanced Deepfake Detection Using Frequency Domain Upsampling <i>Mamadou Dian Bah and Mohamed Dahmane</i>	796
SMART-RD: Towards a Risk Assessment Framework for Autonomous Railway Driving <i>Justin Bescop, Nicolas Goeman, Amel Aissaoui, Benjamin Allaert and Jean-Philippe Vandeborre</i>	804
Image Augmentation for Object Detection and Segmentation with Diffusion Models <i>Leon Useinov, Valeria Efimova and Sergey Muravyov</i>	812
Enhancement-Driven Pretraining for Robust Fingerprint Representation Learning <i>Ekta Gavas, Kaustubh Olpadkar and Anoop Namboodiri</i>	821
Variational Autoencoders for Pedestrian Synthetic Data Augmentation of Existing Datasets: A Preliminary Investigation <i>Ivan Nikolov</i>	829
Unsupervised Domain Adaptation for Human Pose Action Recognition <i>Mattias Billast, Tom De Schepper and Kevin Mets</i>	837
CL-FedFR: Curriculum Learning for Federated Face Recognition <i>Devilliers Caleb Dube, Çiğdem Eroğlu Erdem and Ömer Korçak</i>	845
Vehicle Pose Estimation: Exploring Angular Representations <i>Ivan Orlov, Marco Buzzelli and Raimondo Schettini</i>	853

Exploring Unsupervised Domain Adaptation Approaches for Water Parameters Estimation from Satellite Images <i>Mauren Louise Sguario Coelho de Andrade, Anderson Paulino Souza, Bruno Oliveira, Maria Clara Starling, Camila Costa Amorim and Jefersson A. dos Santos</i>	861
M&M: Multimodal-Multitask Model Integrating Audiovisual Cues in Cognitive Load Assessment <i>Long Nguyen-Phuoc, Renald Gaboriau, Dimitri Delacroix and Laurent Navarro</i>	869
AUTHOR INDEX	877