

20th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2025)

Volume 2: VISAPP

Porto, Portugal
26-28 February 2025

Part 1 of 2

Editors:

**Thomas Bashford-Rogers
Daniel Meneveaux
Mehdi Ammi
Mounia Ziat
Stefan Janicke
Helen Purchase**

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

ISBN: 979-8-3313-1868-0

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© (2025) 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

Crowds and Graphics: Beyond Animation and Visual Effects <i>Julien Pettré</i>	5
On the Importance of Visualisation in a Data Driven Society <i>Daniel Archambault</i>	7
Haptic Intelligence <i>Katherine J. Kuchenbecker</i>	11
Lifelong Visual Representation Learning <i>Diane Larlus</i>	13

IMAGE AND VIDEO UNDERSTANDING

FULL PAPERS

Adaptive Out-of-Distribution Detection with Coarse-to-Fine Grained Representation <i>Kohei Fukuda and Hiroaki Aizawa</i>	19
Distortion-Aware Adversarial Attacks on Bounding Boxes of Object Detectors <i>Pham Phuc, Son Vuong, Khang Nguyen and Tuan Dang</i>	27
Pose-Centric Motion Synthesis Through Adaptive Instance Normalization <i>Oliver Hixon-Fisher, Jarek Francik and Dimitrios Makris</i>	39
ConvKAN: Towards Robust, High-Performance and Interpretable Image Classification <i>Achref Ouni, Chafik Samir, Yousef Bouaziz and Anis Fradi</i>	48
Latent Space Characterization of Autoencoder Variants <i>Anika Shrivastava, Renu Rameshan and Samar Agnihotri</i>	59
Beyond Labels: Self-Attention-Driven Semantic Separation Using Principal Component Clustering in Latent Diffusion Models <i>Felix Stillger, Frederik Hasecke, Lukas Hahn and Tobias Meisen</i>	68
Experience Replay and Zero-Shot Clustering for Continual Learning in Diabetic Retinopathy Detection <i>Gussepe Bravo-Rocca, Peini Liu, Jordi Guitart, Ajay Dholakia, David Ellison and Rodrigo M. Carrillo-Larco</i>	81
Detection of Door-Closing Defects by Learning from Physics-Based Simulations <i>Ryoga Takahashi, Yota Yamamoto, Ryosuke Furuta and Yukinobu Taniguchi</i>	93
Leveraging Vision Language Models for Understanding and Detecting Violence in Videos <i>Jose Alejandro Avellaneda Gonzalez, Tetsu Matsukawa and Einoshin Suzuki</i>	99
Adaptive Prompt Tuning: Vision Guided Prompt Tuning with Cross-Attention for Fine-Grained Few-Shot Learning <i>Eric Brouwer, Jan Erik van Woerden, Gertjan Burghouts, Matias Valdenegro-Toro and Marco Zullich</i>	114

MetaToken: Detecting Hallucination in Image Descriptions by Meta Classification <i>Laura Fieback, Jakob Spiegelberg and Hanno Gottschalk</i>	126
ReST: High-Precision Soccer Player Tracking via Motion Vector Segmentation <i>Fahad Majeed, Khaled Ahmed Lutf Al Thelaya, Nauman Ullah Gilal, Kamilla Swart-Arries, Marco Agus and Jens Schneider</i>	138
Transformer or Mamba for Temporal Action Localization? Insights from a Comprehensive Experimental Comparison Study <i>Zejian Zhang, Cristina Palmero and Sergio Escalera</i>	150
DeepSpace: Navigating the Frontier of Deepfake Identification Using Attention-Driven Xception and a Task-Specific Subspace <i>Ayush Roy, Sk Mohiuddin, Maxim Minenko, Dmitrii Kaplun and Ram Sarkar</i>	163
Self-Supervised Iterative Refinement for Anomaly Detection in Industrial Quality Control <i>Muhammad Aqeel, Shakiba Sharifi, Marco Cristani and Francesco Setti</i>	173
VectorWeaver: Transformers-Based Diffusion Model for Vector Graphics Generation <i>Ivan Jarsky, Maxim Kuzin, Valeria Efimova, Viacheslav Shalamov and Andrey Filchenkov</i>	184
Data-Free Dynamic Compression of CNNs for Tractable Efficiency <i>Lukas Meiner, Jens Mehnert and Alexandru Paul Condurache</i>	196
Enhancing 3D Human Pose Estimation: A Novel Post-Processing Method <i>Elham Iravani, Frederik Hasecke, Lukas Hahn and Tobias Meisen</i>	209
Temporally Accurate Events Detection Through Ball Possessor Recognition in Soccer <i>Marc Peral, Guillem Capellera, Antonio Rubio, Luis Ferraz, Francesc Moreno-Noguer and Antonio Agudo</i>	221
Improving Image Classification Tasks Using Fused Embeddings and Multimodal Models <i>Artur A. Oliveira, Mateus Espadoto, Roberto Hirata Jr. and Roberto M. Cesar Jr.</i>	232
Paint Blob Detection and Decoding for Identification of Honey Bees <i>Andrea P. Gómez-Jaime, Luke Meyers, Josué A. Rodríguez-Cordero, José L. Agosto-Rivera, Tugrul Giray and Rémi Mégret</i>	242
SHORT PAPERS	
Vision-Language In-Context Learning Driven Few-Shot Visual Inspection Model <i>Shiryu Ueno, Yoshikazu Hayashi, Shunsuke Nakatsuka, Yusei Yamada, Hiroaki Aizawa and Kunihito Kato</i>	253
Action Tube Generation by Person Query Matching for Spatio-Temporal Action Detection <i>Kazuki Omi, Jion Oshima and Toru Tamaki</i>	261
CodeSCAN: ScreenCast ANalysis for Video Programming Tutorials <i>Alexander Naumann, Felix Hertlein, Jacqueline Höllig, Lucas Cazzonelli and Steffen Thoma</i>	269
Spiideo SoccerNet SynLoc: Single Frame World Coordinate Athlete Detection and Localization with Synthetic Data <i>Håkan Ardö, Mikael Nilsson, Anthony Cioppa, Floriane Magera, Silvio Giancola, Haochen Liu, Bernard Ghanem and Marc Van Droogenbroeck</i>	278

Deep Image Clustering with Model-Agnostic Meta-Learning <i>Kim Bjerger, Paul Bodesheim and Henrik Karstoft</i>	286
Improving Periocular Recognition Accuracy: Opposite Side Learning Suppression and Vertical Image Inversion <i>Masakazu Fujio, Yosuke Kaga and Kenta Takahashi</i>	298
Segment-Level Road Obstacle Detection Using Visual Foundation Model Priors and Likelihood Ratios <i>Youssef Shoeb, Nazir Nayal, Azarm Nowzad, Fatma Güney and Hanno Gottschalk</i>	306
Classifier Ensemble for Efficient Uncertainty Calibration of Deep Neural Networks for Image Classification <i>Michael Schulze, Nikolas Ebert, Laurenz Reichardt and Oliver Wasenmüller</i>	316
Conditioned Generative AI for Synthetic Training of 6D Object Pose Detection <i>Mathijs Lens, Aaron Van Campenhout and Toon Goedemé</i>	324
Deep Local Feature Matching Image Anomaly Detection with Patch Adaptive Average Pooling Technique <i>Afshin Dini and Esa Rahtu</i>	332
CTypeClust: Confidence-Aware Typical Clustering for Budget-Agnostic Active Learning with Confidence Calibration <i>Takuya Okano, Yohei Minekawa and Miki Hayakawa</i>	340
Neural Network Meta Classifier: Improving the Reliability of Anomaly Segmentation <i>Jurica Runtas and Tomislav Petković</i>	348
New Paths in Document Data Augmentation Using Templates and Language Models <i>Lucas Wojcik, Luiz Coelho, Roger Granada and David Menotti</i>	356
Uncertainty Estimation for Super-Resolution Using ESRGAN <i>Maniraj Sai Adapa, Marco Züllich and Matias Valdenegro-Toro</i>	367
Herbicide Efficacy Prediction Based on Object Segmentation of Glasshouse Imagery <i>Majedaldin Almahasneh, Baihua Li, Haibin Cai, Nasir Rajabi, Laura Davies and Qinggang Meng</i>	375
Inductive Self-Supervised Dimensionality Reduction for Image Retrieval <i>Deryk Willyan Biotto, Guilherme Henrique Jardim, Vinicius Atsushi Sato Kawai, Bionda Rozin, Denis Henrique Pinheiro Salvadeo and Daniel Carlos Guimarães Pedronette</i>	383
A Method for Detecting Hands Moving Objects from Videos <i>Rikuto Konishi, Toru Abe and Takuo Suganuma</i>	392
Rescuing Easy Samples in Self-Supervised Pretraining <i>Qin Wang, Kai Krajsek and Hanno Scharf</i>	400
Knowledge Amalgamation for Single-Shot Context-Aware Emotion Recognition <i>Tristan Cladière, Olivier Alata, Christophe Ducottet, Hubert Konik and Anne-Claire Legrand</i>	410
Handling Drift in Industrial Defect Detection Through MMD-Based Domain Adaptation <i>Xuban Barberena, Fátima A. Saiz and Iñigo Barandiaran</i>	420
Beyond Data Augmentations: Generalization Abilities of Few-Shot Segmentation Models <i>Muhammad Ahsan, Guy Ben-Yosef and Gemma Roig</i>	430

Membership Inference Attacks for Face Images Against Fine-Tuned Latent Diffusion Models <i>Lauritz Christian Holme, Anton Mosquera Storgaard and Siavash Arjomand Bigdeli</i>	439
Minimizing Number of Distinct Poses for Pose-Invariant Face Recognition <i>Carter Ung, Pranav Mantini and Shishir K. Shah</i>	447
VLLM Guided Human-Like Guidance Navigation Generation <i>Masaki Nambata, Tsubasa Hirakawa, Takayoshi Yamashita, Hirobobu Fujiyoshi, Takehito Teraguchi, Shota Okubo and Takuya Nanri</i>	456
CLIP-MDGAN: Multi-Discriminator GAN Using CLIP Task Allocation <i>Shonosuke Gonda, Fumihiko Sakaue and Jun Sato</i>	464
Simultaneous Estimation of Driving Intentions for Multiple Vehicles Using Video Transformer <i>Junya Isogawa, Fumihiko Sakaue and Jun Sato</i>	471
Human Pose Estimation from an Extremely Low-Resolution Image Sequence by Pose Transition Embedding Network <i>Yasutomo Kawanishi, Hitoshi Nishimura and Hiroshi Murase</i>	478
Multi-Scale Foreground-Background Confidence for Out-of-Distribution Segmentation <i>Samuel Marschall and Kira Maag</i>	486
Accuracy Improvement of Neuron Concept Discovery Using CLIP with Grad-CAM-Based Attention Regions <i>Takahiro Sannomiya and Kazuhiro Hotta</i>	497
Expanding Domain Coverage in Injection Molding Quality Inspection with Physically-Based Synthetic Data <i>Dominik Schraml and Gunther Notni</i>	503
Neighbor Embedding Projection and Graph Convolutional Networks for Image Classification <i>Gustavo Rosseto Leticio, Vinicius Atsushi Sato Kawai, Lucas Pascotti Valem and Daniel Carlos Guimarães Pedronette</i>	511
Graph Convolutional Networks and Particle Competition and Cooperation for Semi-Supervised Learning <i>Gustavo Rosseto Leticio, Matheus Henrique Jacob dos Santos, Lucas Pascotti Valem, Vinicius Atsushi Sato Kawai, Fabricio Aparecido Breve and Daniel Carlos Guimarães Pedronette</i>	519
Exploration and Validation of Specialized Loss Functions for Generative Visual-Thermal Image Domain Transfer <i>Simon Fischer, Benedikt Kottler, Eva Strauß and Dimitri Bulatov</i>	527
Semi-Supervised Anomaly Detection in Skin Lesion Images <i>Alina Burgert, Babette Dellen, Uwe Jaekel and Dietrich Paulus</i>	535
Automatic Detection of the Driver Distractions Based on the Analysis of Face Videos <i>Artur Urzędowski and Kazimierz Choroś</i>	542

MOTION, TRACKING, AND 3D VISION

FULL PAPERS

HandMvNet: Real-Time 3D Hand Pose Estimation Using Multi-View Cross-Attention Fusion <i>Muhammad Asad Ali, Nadia Robertini and Didier Stricker</i>	555
MuSt-NeRF: A Multi-Stage NeRF Pipeline to Enhance Novel View Synthesis <i>Sudarshan Raghavan Iyengar, Subash Sharma and Patrick Vandewalle</i>	563
Urban Re-Identification: Fusing Local and Global Features with Residual Masked Maps for Enhanced Vehicle Monitoring in Small Datasets <i>William A. Ramirez, Cesar A. Sierra Franco, Thiago R. da Motta and Alberto Raposo</i>	574
2D Motion Generation Using Joint Spatial Information with 2CM-GPT <i>Ryota Inoue, Tsubasa Hirakawa, Takayoshi Yamashita and Hironobu Fujiyoshi</i>	582
Segmentation-Guided Neural Radiance Fields for Novel Street View Synthesis <i>Yizhou Li, Yusuke Monno, Masatoshi Okutomi, Yuuichi Tanaka, Seiichi Kataoka and Teruaki Kosiba</i>	591
ConMax3D: Frame Selection for 3D Reconstruction Through Concept Maximization <i>Akash Malhotra, Nacéra Seghouani, Gilbert Badaro and Christophe Blaya</i>	598
Improving Adaptive Density Control for 3D Gaussian Splatting <i>Glenn Grubert, Florian Barthel, Anna Hilsmann and Peter Eisert</i>	610
Sensor Calibration and Data Analysis of the MuFoRa Dataset <i>Valentino Behret, Regina Kushtanova, Islam Fadl, Simon Weber, Thomas Helmer and Frank Palme</i>	622
Uncertainty and Feature-Based Weighted Loss for 3D Wheat Part Segmentation <i>R. Reena, John H. Doonan, Kevin Williams, Fiona M. K. Corke, Huaizhong Zhang and Yonghuai Liu</i>	632

SHORT PAPERS

D-PLS: Decoupled Semantic Segmentation for 4D-Panoptic-LiDAR-Segmentation <i>Maik Steinhauser, Laurenz Reichardt, Nikolas Ebert and Oliver Wasenmüller</i>	645
Adaptable Distributed Vision System for Robot Manipulation Tasks <i>Marko Pavlic and Darius Burschka</i>	651
Real-Time Kinematic Positioning and Optical See-Through Head-Mounted Display for Outdoor Tracking: Hybrid System and Preliminary Assessment <i>Muhannad Ismael and Maël Cornil</i>	659
Noisemaker 3D: Comprehensive Framework for Mesh Noise Generation <i>Vladimir Mashurov, Vasilii Latonov, Anastasia Martynova and Natalia Semenova</i>	667
Evaluating Homography Error for Accurate Multi-Camera Multi-Object Tracking of Dairy Cows <i>Shunpei Aou, Yota Yamamoto, Kazuaki Nakamura and Yukinobu Taniguchi</i>	675
FiDaSS: A Novel Dataset for Firearm Threat Detection in Real-World Scenes <i>Murilo S. Regio and Isabel H. Manssour</i>	683
Comparative Analysis of Deep Learning-Based Multi-Object Tracking Approaches Applied to Sports User-Generated Videos <i>Elton Alencar, Larissa Pessoa, Fernanda Costa, Guilherme Souza and Rosiane de Freitas</i>	691

Learning Neural Velocity Fields from Dynamic 3D Scenes via Edge-Aware Ray Sampling <i>Sota Ito, Yoshikazu Hayashi, Hiroaki Aizawa and Kunihito Kato</i>	699
3DSES: An Indoor Lidar Point Cloud Segmentation Dataset with Real and Pseudo-Labels from a 3D Model <i>Maxime Mérizette, Nicolas Audebert, Pierre Kervella and Jérôme Verdun</i>	707
MAESTRO: A Full Point Cloud Approach for 3D Anomaly Detection Based on Reconstruction <i>Remi Lhoste, Antoine Vacavant and Damien Delhay</i>	717
Improving Geometric Consistency for 360-Degree Neural Radiance Fields in Indoor Scenarios <i>Iryna Repinetska, Anna Hilsmann and Peter Eisert</i>	725
Efficient 3D Human Pose and Shape Estimation Using Group-Mix Attention in Transformer Models <i>Yushan Wang, Shuhei Tarashima and Norio Tagawa</i>	735
Leveraging Unreal Engine for UAV Object Tracking: The AirTrackSynth Synthetic Dataset <i>Mingyang Zhang, Kristof Van Beeck and Toon Goedemé</i>	743
Recovery of Detailed Posture and Shape from Motion Video Images by Deforming SMPL <i>Yumi Ando, Fumihiko Sakaue and Jun Sato</i>	751
Deep Learning-Powered Visual SLAM Aimed at Assisting Visually Impaired Navigation <i>Marziyeh Bamdad, Hans-Peter Hutter and Alireza Darvishy</i>	758
Benchmarking Neural Rendering Approaches for 3D Reconstruction of Underwater Environments <i>Salvatore Mario Carota, Alessandro Privitera, Daniele Di Mauro, Antonino Furnari, Giovanni Maria Farinella and Francesco Ragusa</i>	766
An Event Camera Simulator for Arbitrary Viewpoints Based on Neural Radiance Fields <i>Diego Hernández Rodríguez, Motoharu Sonogashira, Kazuya Kitano, Yuki Fujimura, Takuya Funatomi, Yasuhiro Mukaigawa and Yasutomo Kawanishi</i>	774
A Computer Vision Approach to Counting Farmed Fish in Flowing Water <i>Masanori Nishiguchi, Hitoshi Habe, Koji Abe, Masayuki Otani and Nobukazu Iguchi</i>	781
Shape from Mirrored Polarimetric Light Field <i>Shunsuke Nakagawa, Takahiro Okabe and Ryo Kawahara</i>	790
MOBILE, EGOCENTRIC, AND ROBOTIC VISION	
FULL PAPERS	
PrIcosa: High-Precision 3D Camera Calibration with Non-Overlapping Field of Views <i>Oguz Kedilioglu, Tasnim Tabassum Nova, Martin Landesberger, Lijiu Wang, Michael Hofmann, Jörg Franke and Sebastian Reitelshöfer</i>	801
Fine-Grained Self-Localization from Coarse Egocentric Topological Maps <i>Daiki Iwata, Kanji Tanaka, Mitsuki Yoshida, Ryogo Yamamoto, Yuudai Morishita and Tomoe Hiroki</i>	810
GIFF: Graph Iterative Attention Based Feature Fusion for Collaborative Perception <i>Ahmed N. Ahmed, Siegfried Mercelis and Ali Anwar</i>	820

SHORT PAPERS

SSGA: Synthetic Scene Graph Augmentation via Multiple Pipeline Variants <i>Kenta Tsukahara, Ryogo Yamamoto, Kanji Tanaka and Tomoe Hiroki</i>	833
Low Latency Pedestrian Detection Based on Dynamic Vision Sensor and RGB Camera Fusion <i>Bingyu Huang, Gianni Allebosch, Peter Veelaert, Tim Willems, Wilfried Philips and Jan Aelterman</i>	841
Automated Individualization of Object Detectors for the Semantic Environment Perception of Mobile Robots <i>Christian Hofmann, Christopher May, Patrick Ziegler, Iliya Ghotbiravandi, Jörg Franke and Sebastian Reitelshöfer</i>	851
Online Detection of End of Take and Release Actions from Egocentric Videos <i>Alessandro Sebastiano Catinello, Giovanni Maria Farinella and Antonino Furnari</i>	863
Robotic Visual Attention Architecture for ADAS in Critical Embedded Systems for Smart Vehicles <i>Diego Renan Bruno, William D'Abruzzo Martins, Rafael Alceste Berri and Fernando Santos Osório</i>	871

APPLICATIONS AND SERVICES

FULL PAPER

Defense Against Model Inversion Attacks Using a Dummy Recognition Model Trained with Synthetic Samples <i>Yuta Kotsuji and Kazuaki Nakamura</i>	883
--	-----

SHORT PAPERS

Optimum-Path Forest Ensembles to Estimate the Internal Decay in Urban Trees <i>Giovani Candido, Luis Henrique Morelli, Danilo Samuel Jodas, Giuliana Del Nero Velasco, Reinaldo Araújo de Lima, Kelton Augusto Pontara da Costa and João Paulo Papa</i>	895
Coloring 3D Avatars with Single-Image <i>Pin-Yuan Yang, Yu-Shan Deng, Chieh-Shan Lin, An-Chun Luo and Shih-Chieh Chang</i>	903
Internal State Estimation Based on Facial Images with Individual Feature Separation and Mixup Augmentation <i>Ayaka Asaeda and Noriko Takemura</i>	911
Disease Estimation Using Gait Videos by Separating Individual Features Based on Disentangled Representation Learning <i>Shiori Furukawa and Noriko Takemura</i>	919
Towards a Dataset for Paleographic Details in Historical Torah Scrolls <i>Laura Frank, Germaine Götzelmann and Danah Tonne</i>	926
Efficient CNN-Based System for Automated Beetle Elytra Coordinates Prediction <i>Hojin Yoo, Dhanyapriya Somasundaram and Hyunju Oh</i>	934
Effectiveness of Cross-Model Learning Through View-Model Ensemble on Detection of Spatiotemporal EEG Patterns <i>Ömer Muhammet Soysal, Iphy Emeka Kelvin and Muhammed Esad Oztemel</i>	942
A Multimodal Approach to Research Paper Summarization <i>Pranav Bookanakere, Syeda Saniya, Syed Munzer Nouman, S. Pramath and Jayashree Rangareddy</i>	950

Differential Diagnosis of Brain Diseases Using Ensemble Learning and Explainable AI <i>Nighat Bibi, Kathleen M. Curran and Jane Courtney</i>	958
Leveraging Affordable Solutions for Stereo Video Capture in Virtual Reality Applications <i>Leina Yoshida, Gustavo Domingues, Fabiana Peres, Claudio R. Mauricio and João M. Teixeira</i>	965
Sleep-Stage Efficient Classification Using a Lightweight Self-Supervised Model <i>Eldiane Borges dos Santos Durães and João Batista Florindo</i>	972
DeepCellCount: Cell Counting Using Two-Step Deep Learning <i>Sara Tesfamariam, Isah A. Lawal, Arda Durmaz and Jacob G. Scott</i>	980
Towards Safe Self-Stimulatory Behaviors in Autistic Children: HarmAlert4AutisticChildren (HA4AC) <i>Aleenah Khan and Hassan Foroosh</i>	986
AUTHOR INDEX	