

PROCEEDINGS OF SPIE

Seventeenth International Conference on Machine Vision (ICMV 2024)

Wolfgang Osten
Editor

10–13 October 2024
Edinburg, United Kingdom

Organized by
Russian Academy of Sciences (Russian Federation)
University of Stuttgart (Germany)

Sponsored by
Russian Academy of Sciences (Russian Federation)
University of Stuttgart (Germany)

Published by
SPIE

Volume 13517

Proceedings of SPIE 0277-786X, V. 13517

SPIE is an international society advancing an interdisciplinary approach to the science and application of light.

The papers in this volume were part of the technical conference cited on the cover and title page. Papers were selected and subject to review by the editors and conference program committee. Some conference presentations may not be available for publication. Additional papers and presentation recordings may be available online in the SPIE Digital Library at SPIDigitalLibrary.org.

The papers reflect the work and thoughts of the authors and are published herein as submitted. The publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Please use the following format to cite material from these proceedings:

Author(s), "Title of Paper," in *Seventeenth International Conference on Machine Vision (ICMV 2024)*, edited by Wolfgang Osten, Proc. of SPIE 13517, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9781510688278

ISBN: 9781510688285 (electronic)

Published by

SPIE

P.O. Box 10, Bellingham, Washington 98227-0010 USA

Telephone +1 360 676 3290 (Pacific Time)

SPIE.org

Copyright © 2025 Society of Photo-Optical Instrumentation Engineers (SPIE).

Copying of material in this book for internal or personal use, or for the internal or personal use of specific clients, beyond the fair use provisions granted by the U.S. Copyright Law is authorized by SPIE subject to payment of fees. To obtain permission to use and share articles in this volume, visit Copyright Clearance Center at copyright.com. Other copying for republication, resale, advertising or promotion, or any form of systematic or multiple reproduction of any material in this book is prohibited except with permission in writing from the publisher.

Printed in the United States of America by Curran Associates, Inc., under license from SPIE.

Publication of record for individual papers is online in the SPIE Digital Library.

**SPIE. DIGITAL
LIBRARY**

SPIDigitalLibrary.org

Paper Numbering: A unique citation identifier (CID) number is assigned to each article in the Proceedings of SPIE at the time of publication. Utilization of CIDs allows articles to be fully citable as soon as they are published online, and connects the same identifier to all online and print versions of the publication. SPIE uses a seven-digit CID article numbering system structured as follows:

- The first five digits correspond to the SPIE volume number.
- The last two digits indicate publication order within the volume using a Base 36 numbering system employing both numerals and letters. These two-number sets start with 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 0A, 0B ... 0Z, followed by 10-1Z, 20-2Z, etc. The CID Number appears on each page of the manuscript.

Contents

vii	<i>Conference Committees</i>
xi	<i>Introduction</i>

IMAGE SEGMENTATION AND COMPUTING

13517 02	A deep learning-based retinal disease analysis framework for optical coherence tomographic image generation, segmentation, and classification [13517-55]
13517 03	Barrel instance segmentation and grasping for (semi-)autonomous excavators [13517-24]
13517 04	Impact of synthetic images in the training of neural networks for airborne vessel segmentation [13517-50]
13517 05	Multimodality fusion for enhanced driving scene semantic segmentation [13517-19]

COMPUTER VISION AND IMAGE ANALYSIS

13517 06	Dataset and feature point-based matching algorithm for satellite and aerial remotely sensed images [13517-25]
13517 07	VGG16, ResNet50, and Xception feature extractors combined with various classifiers for autism spectrum disorder classification using fMRI data [13517-46]
13517 08	Infrared domain adaptation with zero-shot quantization [13517-12]
13517 09	TSQ-2024: a categorized dataset of 2D LiDAR images of moving dump trucks in various environment conditions [13517-36]
13517 0A	CS-MAT: cross-shaped window mask-aware transformer for large mask image inpainting [13517-40]
13517 0B	A new method for manually calculating camera pose [13517-13]
13517 0C	Improving pain classification using spatio-temporal deep learning approaches with facial expressions [13517-52]

- 13517 OD **Lightweight adaptive learning algorithm for energy-latency tradeoff in IoMT-enabled edge computing** [13517-56]
- 13517 OE **Diagnosing ADHD with 1D-CNN: efficient analysis of fMRI data** [13517-59]
- 13517 OF **SVI: stable video index for video stabilization quality assessment** [13517-61]

IMAGE-BASED INTELLIGENT DETECTION TECHNOLOGY AND APPLICATION

- 13517 OG **A black-box technique to mitigate gender bias for generative AI** [13517-35]
- 13517 OH **Electoral symbols and vote detection in paper ballots: a case from Nepal's election** [13517-1]
- 13517 OI **X-ray anomaly detection in industrial pipelines** [13517-3]
- 13517 OJ **A light perspective for 3D object detection** [13517-5]
- 13517 OK **Diffusion model uniform manifold filtering for classification of small datasets with underrepresented classes: application to chromosomal aberration microscopy detection** [13517-7]
- 13517 OL **Convolutional neural network and k-nearest neighbors classifier for enhanced detection of multiple sclerosis** [13517-10]
- 13517 OM **MTFL: multi-timescale feature learning for weakly-supervised anomaly detection in surveillance videos** [13517-14]
- 13517 ON **Soiling detection for advanced driver assistance systems** [13517-39]
- 13517 OO **Approach to lightweighting backbone network models through quantization and Bayesian optimization** [13517-49]
- 13517 OP **Single object spectrum training for 2D NMR signals detection** [13517-64]
- 13517 OQ **Detecting computer-generated images by using only real images** [13517-18]
- 13517 OR **An improved 3D skeletons UP-Fall dataset: enhancing data quality for efficient impact fall detection** [13517-31]

MACHINE VISION AND MEASUREMENT BASED ON MACHINE LEARNING

- 13517 OS **EEPPR: event-based estimation of periodic phenomena rate using correlation in 3D** [13517-4]
- 13517 OT **Graph convolutional neural networks for skull stripping** [13517-29]
- 13517 OU **Blur patch classification approach to single-image depth estimation** [13517-32]
- 13517 OV **Ablation study for multicamera vehicle tracking using CityFlow dataset** [13517-38]
- 13517 OW **Enriching visual information in images through fusion with thermal imaging in low-light surveillance scenes** [13517-2]
- 13517 OX **High-throughput median filtering for large kernel sizes on CUDA** [13517-28]

DETECTION AND CLASSIFICATION TECHNOLOGY OF FOOD AND PLANT DISEASES BASED ON MACHINE VISION

- 13517 OY **Leveraging stable-diffusion generated images to enhance vision transformers for plant disease classification** [13517-69]
- 13517 OZ **Tomato disease detection to reduce food waste** [13517-6]
- 13517 IO **Image-based quantification of scale loss in fish using machine learning and computer vision** [13517-37]
- 13517 I1 **KonvLiNA: integrating Kolmogorov-Arnold Network with linear Nyström attention for feature fusion in crop field detection** [13517-16]
- 13517 I2 **iiADET-tiny: enhanced real-time rice leaf disease detection on low-power embedded systems** [13517-17]
- 13517 I3 **Protecting beehive health: varroa mite detection using gen-AI** [13517-48]
- 13517 I4 **High-resolution leaf disease classification using deep learning advancing agricultural technology through innovative approaches** [13517-54]

ADVERSARIAL MACHINE LEARNING IN VISION, SPEECH, AND TEXT

- 13517 I5 **Generating synthetic sign language datasets using conditional generative adversarial networks** [13517-33]
- 13517 I6 **Enhancing visual odometry performance through fine-tuning: a case study with TartanVO** [13517-67]

- 13517 17 **Multimodal approach for imbalanced document classification** [13517-21]
- 13517 18 **Visible coalitions of neuronal activities in brain-to-text communication via handwriting** [13517-68]
- 13517 19 **Local masking meets progressive freezing: crafting efficient vision transformers for self-supervised learning** [13517-30]
- 13517 1A **Annotation tools for computer vision tasks** [13517-11]
- 13517 1B **Trans-APL: transformer model for audio and prior landmark fusion for talking landmark generation** [13517-15]
- 13517 1C **Using autoencoders to improve classification of online Arabic handwriting: addressing the homogeneity problem in data** [13517-53]

INTELLIGENT IMAGE RECOGNITION AND APPLICATION TECHNOLOGY

- 13517 1D **E-Gait: enhanced efficient graph convolution network for gait recognition** [13517-9]
- 13517 1E **Optical braille recognition of uncontracted unified English and Filipino braille** [13517-20]
- 13517 1F **3D action recognition with motion encoded RGB images using deep learning** [13517-45]
- 13517 1G **A survey on the online Arabic handwriting recognition: challenges, datasets, and future directions** [13517-51]
- 13517 1H **MobileViTv2-MLP: Arabic traffic sign recognition with enhanced lightweight vision transformers** [13517-22]
- 13517 1I **Beehive state and events recognition through sound analysis using TinyML** [13517-47]
- 13517 1J **Black-box adversarial defense for enhancing robustness in speaker recognition systems with multimodel consensus** [13517-58]