

PROCEEDINGS OF SPIE

Eighteenth International Conference on Machine Vision (ICMV 2025)

Wolfgang Osten
Editor

19–22 October 2025
Paris, France

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 14114

Proceedings of SPIE 0277-786X, V. 14114

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 *Eighteenth International Conference on Machine Vision (ICMV 2025)*, edited by Wolfgang Osten, Proc. of SPIE 14114, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9798902321873

ISBN: 9798902321880 (electronic)

Published by

SPIE

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

Telephone +1 360 676 3290 (Pacific Time)

SPIE.org

Copyright © 2026 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

- ix *Conference Committees*
- xv *Introduction*

IMAGE RECOGNITION AND PATTERN RECOGNITION

- 14114 02 **Spectral band attribution in historical ink and substrate recognition** [14114-45]
- 14114 03 **Real-time feature recognition of cantaloupe flowers and buds in greenhouse based on deep learning for pollination robot** [14114-47]
- 14114 04 **Real-time OCR-based grocery product recognition with orientation alignment and embedding-driven classification** [14114-100]
- 14114 05 **Precise recognition and localization of drill pipes based on SBN-YOLO and stereo vision** [14114-39]
- 14114 06 **Real-time automatic water meters recognition on edge devices** [14114-105]
- 14114 07 **A novel composite metric loss for offline handwritten Chinese character recognition with CNNs** [14114-68]
- 14114 08 **Hybrid CNN-GNN for offline handwritten math expression recognition** [14114-63]
- 14114 09 **Orthogonality-aware projection-based feature enhancement for fine-grained facial expression recognition in the wild** [14114-8]
- 14114 0A **WholeBodyPose: a unified end-to-end framework for sign language recognition and pose-based training data** [14114-49]
- 14114 0B **Asthma identification through multimodal data fusion: towards reliable clinical decision support** [14114-77]
- 14114 0C **CFE-FER: channel-guided feature enhancement for robust facial expression recognition** [14114-7]
- 14114 0D **Enhancing Arabic traffic sign recognition through synthetic data generation using diffusion models** [14114-114]

IMAGE SEGMENTATION

- 14114 0E **Automatic detection and segmentation of squamous cervical cells in Papanicolaou smear-images** [14114-99]
- 14114 0F **Foundation-guided active learning for efficient ultrasound image segmentation** [14114-20]

- 14114 0G **Is negative correlation learning a solution for cross domain colon polyp segmentation?** [14114-88]
- 14114 0H **Shapedistill: shape-constrained knowledge distillation for medical segmentation** [14114-2]
- 14114 0I **Colorectal cancer segmentation with adaptive augmentation and multiresolution ensemble models** [14114-73]
- 14114 0J **Self-distillation with edge-aware pseudo-masks for skin lesion segmentation** [14114-38]
- 14114 0K **FFT-based appearance adaptation for the semantic segmentation of small-scale aerial image datasets** [14114-32]
- 14114 0L **IIRDoc-Net: efficient ID segmentation with learnable IIR filters for mobile and embedded platforms** [14114-69]

IMAGE CLASSIFICATION AND ALGORITHMS

- 14114 0M **Video game-driven EEG signal classification for rehabilitation applications using LSTM networks** [14114-46]
- 14114 0N **Fine-tuning vision-language models enables robust classification of naturally imbalanced camera trap images** [14114-102]
- 14114 0O **Deep-learning-driven classification of bone fractures for emergency and disaster response applications** [14114-50]
- 14114 0P **Olive leaf diseases classification based on feature fusion using vision transformer and CNNs** [14114-85]
- 14114 0Q **UNet-HoVerGNN: structured graph integration into HoVerNet for enhanced nuclei segmentation and classification** [14114-37]
- 14114 0R **Hy-Facial: hybrid feature extraction by dimensionality reduction methods for enhanced facial expression classification** [14114-43]
- 14114 0S **Shallow and explainable LiDAR-based weather classifier neural network for automated driving** [14114-80]
- 14114 0T **Domain-specific convolutional neural network layers for galaxy classification** [14114-5]
- 14114 0U **Adaptive augmentation policy optimization with LLM feedback** [14114-72]
- 14114 0V **VNIR hyperspectral imaging for Nordic winter road surface classification in autonomous driving applications** [14114-34]
- 14114 0W **Multiview clustering via self-aware contrastive attention fusion** [14114-66]
- 14114 0X **Joint MK classification of stellar spectra with distance-aware CNNs** [14114-33]

- 14114 0Y **From teacher to student: achieving high performance with compact models for waste classification** [14114-86]
- 14114 0Z **A hybrid framework for classifying interior design with attention transformer** [14114-113]

MACHINE VISION AND VISUALIZATION

- 14114 10 **Sub-pixel concentricity estimation algorithm for industrial machine vision with an accompanying dataset** [14114-1]
- 14114 11 **TFM-VO: transformer-fused multiscale visual odometry** [14114-27]
- 14114 12 **Inter-UAV distance estimation using a variable baseline stereo vision system** [14114-51]
- 14114 13 **Fast- and memory-efficient convolution on ARM for computer vision neural networks** [14114-57]
- 14114 14 **SAM-PMIL: smoothness-consistent adversarial masking proposal-based multiple instance learning for weakly supervised temporal action localization** [14114-108]
- 14114 15 **A T-spherical fuzzy MEREC-CoCoSo framework for evaluating machine vision systems in manufacturing quality control** [14114-52]
- 14114 16 **Quantum-assisted visual quality assessment of pomegranates using hybrid transfer learning** [14114-61]
- 14114 17 **Enhancing one-step diffusion models using GANs with application to mental health mindfulness** [14114-97]

INTELLIGENT IMAGE DETECTION MODELS AND ALGORITHMS

- 14114 18 **Channel-attentive transformer-based multimodal semantic segmentation model for early detection of wheat yellow rust disease** [14114-19]
- 14114 19 **Towards robust person detection: evaluating model performance in objectifying and visually constrained scenes** [14114-84]
- 14114 1A **Wind turbine feature detection using deep learning and synthetic data** [14114-112]
- 14114 1B **Morphological real-time defense against physical attacks on convolutional neural networks** [14114-29]
- 14114 1C **ArrowFormer: detecting painted symbols using DEViT** [14114-40]
- 14114 1D **Evaluation of pre-trained convolutional networks for beverage cans defect inspection** [14114-111]
- 14114 1E **MEFA-MS: attention-based U-Net for pedestrian and vehicle detection** [14114-24]

- 14114 1F **Depth-guided self-supervised human keypoint detection via cross-modal distillation** [14114-87]
- 14114 1G **Dataset curation for a domain-specific people detection system** [14114-11]
- 14114 1H **Human perception for print defect detection** [14114-81]
- 14114 1I **Accelerated reinforcement learning-based pruning strategy search for cross-domain generalization** [14114-59]
- 14114 1J **Real-time fish detection and counting system using deep learning computer vision** [14114-48]
- 14114 1K **FusionViT: hierarchical 3D object detection via Lidar-camera vision transformer fusion** [14114-78]
- 14114 1L **Patch-based localized artifact enhancement for deepfake detection** [14114-55]
- 14114 1M **A CNN-LSTM hybrid framework with attention mechanism for robust tomato disease detection** [14114-92]
- 14114 1N **RGB-thermal drone imaging for semantic segmentation and anomaly detection in photovoltaic panels** [14114-75]

MEDICAL IMAGE DETECTION AND AUTOMATIC DIAGNOSIS

- 14114 1O **SLOTMFound: foundation-based diagnosis of multiple sclerosis using retinal SLO imaging and OCT thickness-maps** [14114-18]
- 14114 1P **Boosting high-resolution 3D point cloud anomaly detection with geometric constraints** [14114-62]
- 14114 1Q **Hyperspectral imaging-based framework for automated detection of dysplastic melanocytes in dermatopathology** [14114-21]
- 14114 1R **A CNN-transformer hybrid method for kidney disease detection** [14114-94]
- 14114 1S **An efficient pneumonia detection in chest x-ray images using EfficientNetB4 with attention mechanism** [14114-35]
- 14114 1T **Madvit: multiview Alzheimer's disease detection based on vision transformer networks** [14114-41]
- 14114 1U **Spectro-temporal audio fusion for early Parkinson's detection** [14114-74]
- 14114 1V **GOOSE-Net: Gompertz-augmented overlapping output-based segmentation enhancement for detecting colon polyp** [14114-89]
- 14114 1W **Malaria detection system with machine learning and deep learning techniques** [14114-58]

- 14114 1X **Adaptive neuroscience: leveraging liquid neural networks for early Parkinson's disease detection** [14114-65]

COMPUTER-AIDED GRAPHICS AND INTELLIGENT IMAGE PROCESSING TECHNOLOGY

- 14114 1Y **CD-HMA: degradation-aware hybrid multiaxis attention network for blind image super-resolution** [14114-76]
- 14114 1Z **Five dimensional interferenceless coded aperture correlation holography with a vortex filter** [14114-64]
- 14114 20 **Advanced scattering media removal by modified ARMS and restoration of color information** [14114-4]
- 14114 21 **LH-MemUDA: low-high resolution memory black-box unsupervised domain adaptation** [14114-54]
- 14114 22 **Towards high-quality 3D scans of orchards for automated pruning** [14114-79]
- 14114 23 **Quality assurance with visual inspection in sheet metal forming process for automotive industry** [14114-106]
- 14114 24 **InnoGuard: secure authorship verification using ECC-encoded watermarks and randomized embedding for AIGC-resilient images** [14114-107]
- 14114 25 **UDRSim: an underwater dynamic real-time multimodal imaging simulator** [14114-28]
- 14114 26 **PSinSR: perceptual single step for diffusion-based image super resolution** [14114-6]
- 14114 27 **Polarization-resolved coded aperture imaging techniques** [14114-91]

INNOVATIVE APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN DIGITAL IMAGE AND SIGNAL SYSTEMS

- 14114 28 **Hierarchical capsule networks with domain-specific pre-training for robust audio-visual speech recognition** [14114-13]
- 14114 29 **Textureless CAD-based domain randomized synthetic data generation for AI model training** [14114-70]
- 14114 2A **Quality-adaptive late fusion for capsule-based audio-visual speech recognition** [14114-14]
- 14114 2B **Fast approximate matrix multiplication for 8-bit neural networks using tree averaging** [14114-36]
- 14114 2C **Improving meta-learned loss function accuracy and convergence speed through normalisation for image classification tasks** [14114-96]

TEXT IMAGE UNDERSTANDING AND IMAGING TECHNOLOGY

- 14114 2D **Improving signature presence detection in identity documents with spatial attention mechanism** [14114-109]
- 14114 2E **Detection of photocopied documents in remote identity verification by presence of the paper frame** [14114-67]
- 14114 2F **Fast and accurate page orientation detector for document analysis systems** [14114-60]
- 14114 2G **SkillFormer: unified multiview video understanding for proficiency estimation** [14114-31]
- 14114 2H **Attribute localization in semi-structured document images applying cozy extractors** [14114-53]

IMAGE MODELS AND TRANSFORM ANALYSIS METHODS

- 14114 2I **Explicit disentanglement through conditional modeling of image properties** [14114-12]
- 14114 2J **Towards higher HRNet accuracy for pose estimation using motion capture-driven annotations** [14114-15]
- 14114 2K **Fine-tuning Gaussian denoisers for fluorescence microscopy image enhancement** [14114-16]
- 14114 2L **A high-accuracy fast Hough transform with linear-log-cubed computational complexity for arbitrary-shaped Images** [14114-30]
- 14114 2M **Generating synthetic datasets to train deep models for counting in large-scale animal groups** [14114-56]
- 14114 2N **Personalized stable diffusion model for olive leaf image generation with disease symptoms** [14114-71]

OPTIMIZED APPLICATIONS AND INTELLIGENT DIAGNOSIS OF VISUAL LANGUAGE MODELS IN INFORMATION MEDICINE

- 14114 2O **Medical image retrieval using multimodal vector similarity** [14114-103]
- 14114 2P **Silhouette extraction from depth sensors for older adult monitoring at nursing homes** [14114-10]
- 14114 2Q **Optimized explainable deep learning model for lung cancer diagnosis from computed tomography images** [14114-101]
- 14114 2R **Single haze removal method using peplography** [14114-9]
- 14114 2S **Low-risk pap smear image filtering method via transfer learning** [14114-25]