

PROCEEDINGS OF SPIE

International Conference on Computer Vision, Robotics, and Automation Engineering (CRAE 2025)

Ning Sun
Ata Jahangir Moshayedi
Editors

27–29 June 2025
Shanghai, China

Organized by
Shanghai Jiao Tong University (China)

Sponsored by
Shandong Normal University (China)
University of Nottingham Ningbo China (China)
University of Electronic Science and Technology of China (China)
AEIC—Academic Exchange Information Centre (China)

Published by
SPIE

Volume 13790

Proceedings of SPIE 0277-786X, V. 13790

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 *International Conference on Computer Vision, Robotics, and Automation Engineering (CRAE 2025)*, edited by Ning Sun, Ata Jahangir Moshayedi, Proc. of SPIE 13790, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9781510694590

ISBN: 9781510694606 (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

xi *Conference Committee*

VISION-GUIDED PERCEPTION AND MEASUREMENT

- 13790 02 **Topics identification based on user discussions in an electric vehicle forum in China**
[13790-64]
- 13790 03 **Clothing violation detection under occlusion conditions** [13790-224]
- 13790 04 **Study on the measurement accuracy of ultrasonic flow meters after ball valves** [13790-110]
- 13790 05 **Coal gangue target detection algorithm based on MBP-YOLOv8** [13790-136]
- 13790 06 **Vision-based pose estimation for cooperative navigation of AUVs** [13790-211]
- 13790 07 **Defect recognition of tunnel pavement based on laser radar point cloud data and 3D reconstruction technology** [13790-95]
- 13790 08 **Spectral prediction of CMYK printed images using stacking ensemble learning** [13790-97]
- 13790 09 **CSI-Net: revisit content style interaction in artistic style transfer** [13790-89]
- 13790 0A **Space and semantics sensitive segmentation using derivative** [13790-62]
- 13790 0B **Feature decoupling and recombination for weakly-supervised semantic segmentation**
[13790-63]
- 13790 0C **Research and analysis of image recognition models based on optical diffraction neural networks** [13790-176]
- 13790 0D **Multiple-level threshold for image partitioning exploiting median greyscale-based 2D entropy in 4-neighbourhood** [13790-111]
- 13790 0E **SN-YOLO: a super neck algorithm based on YOLO11 for traffic object detection** [13790-204]
- 13790 0F **A multitarget multi-camera tracking method for terminal scene** [13790-101]
- 13790 0G **Research on photonic neural network and high-precision image recognition technology**
[13790-131]

- 13790 OH **Research on tobacco leaf mildew detection method based on multichannel imaging system** [13790-120]
- 13790 OI **Research on fish visual recognition technology based on YOLOv5 fishery robot** [13790-134]
- 13790 OJ **WTD-YOLO: an improved wind turbine defect detection algorithm** [13790-36]
- 13790 OK **Research on remote sensing images based on Swin Transformer** [13790-115]
- 13790 OL **The research on a method for detecting the appearance quality of de-stemmed fresh Shiitake mushrooms based on machine vision** [13790-99]
- 13790 OM **Baseline correction of Raman spectra using static dropout triangular deep convolutional network** [13790-114]
- 13790 ON **Infrared and visible image fusion based on cross-modality feature alignment** [13790-213]
- 13790 OO **Image segmentation and recognition method for different stages of *Lyophyllum decastes* drying based on YOLO v8** [13790-162]
- 13790 OP **A classroom check-in algorithm based on face recognition** [13790-59]
- 13790 OQ **An automatic identification method of vital sign parameters based on YOLOv10 and Shape-IoU** [13790-80]
- 13790 OR **Enhanced multibranch Taylor transformer for real-world image de-fogging method** [13790-208]
- 13790 OS **Research on intelligent detection methods for tobacco packaging quality based on 3D point cloud** [13790-202]
- 13790 OT **Saliency-based dual-layer contrast method for infrared small target detection** [13790-157]
- 13790 OU **Adjacent paired LBP and enhanced HOG for improved facial expression recognition** [13790-31]
- 13790 OV **Bibliometric analysis of the application of computer vision in intelligent driving: based on Web of Science core collection data** [13790-45]
- 13790 OW **Design and implementation of a 3D industrial model generator** [13790-61]
- 13790 OX **AttenGaze: an attention-guided gaze segmentation model for efficient human-computer interaction** [13790-161]
- 13790 OY **ResNet-based image recognition model: interpretability and visualization research** [13790-8]

- 13790 0Z **Xception-CM: a deep learning approach for non-invasive skin barrier function assessment using dermoscopic images** [13790-179]
- 13790 10 **ESFormer: edge enhancement SparseViT for accurate liver tumor segmentation** [13790-145]
- 13790 11 **Temporal-aware multimodal event network for dense audio-visual event localization** [13790-128]
- 13790 12 **Research on fire monitoring of Mount Kilimanjaro based on Sentinel-1 images** [13790-25]
- 13790 13 **Multimodal MRI image fusion network for brain tumors based on feature cross enhancement** [13790-46]

ROBOTICS, AUTOMATION, AND CONTROL

- 13790 14 **Small sample reliability modeling of CNC machine tools based on Bayes theory** [13790-98]
- 13790 15 **Dynamic modeling of machine tool spindle system based on substructure synthesis method** [13790-1]
- 13790 16 **Research on process path optimization of defective products in 3D printing based on digital technology** [13790-201]
- 13790 17 **Research on electromagnetic tomography method for imaging reinforcement structures in concrete utility poles based on an improved VGG16 deep learning network** [13790-50]
- 13790 18 **A Stackelberg-guided deep Q-network framework for risk-aware decision-making in autonomous highway driving** [13790-2]
- 13790 19 **The design and experiment of mushroom root cutting robot** [13790-144]
- 13790 1A **Design and research of automatic recognition sole edge polishing system** [13790-6]
- 13790 1B **A heterogeneous population collaborative dual-layer ant colony algorithm for mobile robot path planning** [13790-205]
- 13790 1C **NavigateSense: a multimodal smart blind cane system with deep learning for enhanced environmental perception and mobility assistance** [13790-107]
- 13790 1D **Design and implementation of balance vehicle** [13790-66]
- 13790 1E **Effectiveness assessment of UAV reconnaissance via integrated ANP and ADC-SEA models** [13790-170]
- 13790 1F **Research on intelligent positioning of initial welding point of mobile welding robot on curved structures** [13790-210]

- 13790 1G **Intelligent simulation system and experimental verification of shock absorber based on lightweight** [13790-108]
- 13790 1H **Modeling and simulation of vehicle dynamics based on Simulink for braking conditions** [13790-65]
- 13790 1I **Intelligent algorithm-driven automatic evaluation system for handling stability** [13790-165]
- 13790 1J **Dynamics modeling and integrated control of a spider parallel robot system** [13790-217]
- 13790 1K **G-VOM++: precision off-road obstacle detection via dual-LiDAR fusion and plane calibration** [13790-171]
- 13790 1L **Research on action recognition method based on multimodal fusion of HRNet, IMU, and Kinect** [13790-148]
- 13790 1M **Research on wheel-leg balancing robot based on cascade PID algorithm** [13790-142]
- 13790 1N **Research on enhancing mechanical equipment model quality based on Cesium physically-based rendering (PBR) technology** [13790-173]
- 13790 1O **Multi-region perspective transformation-based measurement of pavement disease size** [13790-41]
- 13790 1P **Precision harvesting for premium tea buds: overcoming recognition and damage-control challenges via multisensor fusion** [13790-146]
- 13790 1Q **A multiobjective giant armadillo algorithm based on two-archive for UAV path planning** [13790-122]
- 13790 1R **ERGTrack: a drone-based method for tracking multiple small objects** [13790-190]
- 13790 1S **Autonomous reconfigurable robotic system** [13790-57]
- 13790 1T **Robust SLAM through adaptive feature detection and multiscale analysis** [13790-133]
- 13790 1U **Method on fisheye image rectification based on distortion feature extraction** [13790-109]
- 13790 1V **Prompt-adaptive continual learning: integrating prompt-based learning with reinforcement learning for robust robotic task adjustment** [13790-29]
- 13790 1W **Incremental stitching for real-time video processing: a comprehensive performance study** [13790-7]
- 13790 1X **Control of moisture content of rewarming equipment based on model predictive theory** [13790-168]
- 13790 1Y **A YOLO-based citrus image recognition and harvesting system** [13790-104]

- 13790 1Z **Robot localization and environment mapping based on SC-LIO-SAM** [13790-127]
- 13790 20 **Multilayer perception-based intelligent feeding algorithm and its deployment in mobile autonomous aquaculture systems** [13790-199]

AI FOR HEALTH AND INFRASTRUCTURE MONITORING

- 13790 21 **Critical information infrastructure security posture forecast** [13790-74]
- 13790 22 **Enhanced association rule mining for gas cylinder accident analysis** [13790-18]
- 13790 23 **Information-based fast drilling technology applied in Shunbei region** [13790-194]
- 13790 24 **Research and implementation of an intelligent diagnosis platform for chronic diseases in internal medicine based on cloud computing** [13790-155]
- 13790 25 **Diagnostic method for main spoke bolt loosening in drum screen based on weighted root mean square error** [13790-21]
- 13790 26 **Optimized pneumonia detection using deep learning: a ResNet50-based model with multiscale efficient spatial attention** [13790-27]
- 13790 27 **Medical image segmentation via a learning structure based on an enhanced transformer and CNN** [13790-28]
- 13790 28 **N2AN-DA: a data augmentation pipeline for limited anomalous data using image-to-image translation** [13790-47]
- 13790 29 **Multi-strategy enhanced privacy protection for multimodal data in federated learning** [13790-30]
- 13790 2A **Enhanced descriptive analysis capabilities of large multimodal models for intelligent substation operation and maintenance** [13790-76]
- 13790 2B **A meteorological disaster prediction model based on U-Net** [13790-94]
- 13790 2C **Construction safety behavior recognition and classification for port engineering: an empirical study based on machine learning** [13790-15]
- 13790 2D **Research on privacy-preserving cybersecurity situation awareness model based on federated learning** [13790-222]
- 13790 2E **YOLO-MEDX: efficient brain tumor detection with YOLO11-EfficientNet** [13790-58]
- 13790 2F **Research on human fall detection based on lightweight YOLO and deployment with PyTorch mobile** [13790-53]

- 13790 2G **Pattern classification based on radial basis function neural network** [13790-121]
- 13790 2H **Deep tobacco leaf grading using ShuffleNet and self-adaptive attention** [13790-69]
- 13790 2I **Gas turbine rotor fault diagnosis method based on FFT-CNN transformer and cross attention mechanism** [13790-152]
- 13790 2J **TriDiffSTG: traffic flow prediction with spatiotemporal graph using triplet attention and diffusion model** [13790-178]
- 13790 2K **AdaVis: adaptive training framework and system optimization for large-scale visual data** [13790-223]
- 13790 2L **ResNet50-based satellite remote sensing application for temporary land use monitoring** [13790-4]
- 13790 2M **Dual decoder multimodal salient object detection** [13790-198]
- 13790 2N **PCB micro-defect detection algorithm based on receptive field attention** [13790-102]
- 13790 2O **A pathological image-text interaction model from the perspective of state space models** [13790-126]
- 13790 2P **Pulmonary nodule detection model based on YOLOv8-MSDA** [13790-216]
- 13790 2Q **Multimodal emotion recognition based on multi-head cross-attention mechanism** [13790-92]
- 13790 2R **Fishery resource intelligent monitoring based on granularity-fused improved YOLOv5** [13790-207]
- 13790 2S **Spatiotemporal attention-driven joint reconstruction of elevator audio and vibration signals via dual-modality fusion** [13790-215]
- 13790 2T **ANormGS: normal prior guided anchored 3D Gaussian splatting for multiview scene reconstruction** [13790-166]
- 13790 2U **GlobalFlow-former: a transformer network based on global context for accurate mainstream line identification in complex river networks** [13790-23]
- 13790 2V **Multitask segmentation of vessels, catheters, and guidewires in medical imaging using a YOLOP-UNet hybrid architecture** [13790-159]
- 13790 2W **SkinSeg-cascade: a two-stage fusion framework for robust face detection via skin-region prior and cascade verification** [13790-33]
- 13790 2X **Analysis and research on traditional Zhuang costume patterns based on an improved MobileNetV3** [13790-172]

- 13790 2Y **A person re-identification system based on Swin-transformer network [13790-197]**
- 13790 2Z **Intelligent diagnostic model for pulmonary disease based on auto-metric graph neural network with chest CT images [13790-32]**
- 13790 30 **Chest CT images analysis based on U-Net architecture and three-dimensional radiomics feature quantification [13790-24]**
- 13790 31 **Automatic tooth segmentation method based on mask scoring RCNN in cone-beam CT images [13790-180]**
- 13790 32 **Cardiovascular disease detection by leveraging semisupervised learning [13790-160]**