PROCEEDINGS OF SPIE

Third International Conference on Image Processing, Object Detection, and Tracking (IPODT 2024)

Bin Liu
Lu Leng
Editors

9–11 August 2024
Nanjing, China

Organized by
Xi'an University of Technology (China)

Published by
SPIE

Volume 13396

Proceedings of SPIE 0277-786X, V. 13396

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 SPIEDigitalLibrary.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 *Third International Conference on Image Processing, Object Detection, and Tracking (IPODT 2024)*, edited by Bin Liu, Lu Leng, Proc. of SPIE 13396, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X
ISSN: 1996-756X (electronic)

ISBN: 9781510685512
ISBN: 9781510685529 (electronic)

Published by
SPIE
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone +1 360 676 3290 (Pacific Time)
SPIE.org
Copyright © 2024 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
SPIEDigitalLibrary.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 *Conference Committee*
ix *Introduction*

MACHINE VISION AND IMAGE PROCESSING TECHNOLOGY

13396 02 **U-Trans Net: a hybrid of CNN and transformer network for liver and liver tumor segmentation** [13396-7]

13396 03 **Enhancing image quality assessment with ResNet50 and global average pooling** [13396-3]

13396 04 **Research on classification and recognition of CFRP cylinder fiber layer fracture based on Wigner-Ville distribution images and HRNet** [13396-29]

13396 05 **3D-CA Net: a 3D-CNN with composite multidimensional attention for the assessment of lung aging** [13396-31]

13396 06 **Construction consumption measurement method based on convolutional cloud image recognition technique** [13396-38]

13396 07 **Study on detecting the influence of sunlight based on infrared thermal-imaging method** [13396-39]

13396 08 **CAM: compound attention module for improved vision models** [13396-23]

13396 09 **Multiview and multidescription video coding based on pattern and prediction vector replication** [13396-20]

13396 0A **Design of a book misplacement detection system based on image recognition** [13396-13]

13396 0B **Automatic detection of ultrasonic water meter LCD pattern elements based on ConvNeXt** [13396-18]

13396 0C **Deep intrinsic image popularity assessment by learning to rank** [13396-26]

13396 0D **A late fusion framework for multivehicle collaborative perception** [13396-14]

13396 0E **Subpixel precision BGA chip localization method based on Lagrange interpolation** [13396-21]

13396 0F **Signal-to-noise ratio analysis and experimental verification of imaging laser detection system** [13396-33]

13396 0G **Comparison of cell image segmentation results based on UNet and transformer** [13396-30]

13396 0H **Deep-learning-based algorithm for classifying pedestrian behavior at crosswalks** [13396-37]

INTELLIGENT TARGET RECOGNITION AND DETECTION TECHNOLOGY

13396 0I **A multiscale balanced aggregation network model for object detection** [13396-41]

13396 0J **Lightweight YOLOv8 for grape leaf lesion detection** [13396-28]

13396 0K **Lightweight satellite recognition algorithm based on improved YOLOv5** [13396-16]

13396 0L **Early prediction method for frost damage and crack growth of expressway pavement based on infrared thermal imaging technology** [13396-36]

13396 0M **Deep learning in food category recognition** [13396-24]

13396 0N **Improved pointpillar point cloud object detection based on channel attention mechanism** [13396-35]

13396 0O **A UAV object tracking algorithm based on *a*-EIOU-NET** [13396-11]

13396 0P **Fault diagnosis of wind turbine gearbox based on wavelet packet denoising and CNN-Swin Transformer-LSTM** [13396-19]

13396 0Q **Detection and feature fusion of trajectory Poisson multi-Bernoulli mixture algorithms for visual multiobject tracking** [13396-40]

13396 0R **Drone-based bistatic SAR-infrared cross-modality vehicle detection and classification** [13396-27]

13396 0S **Research on dangerous driving behavior recognition method based on convolutional neural network** [13396-6]

13396 0T **An improved stamen target detection method based on YOLOv8** [13396-8]

13396 0U **The application and research of improved YOLOv10 algorithm in smoking behavior detection on campus** [13396-17]

13396 0V **Defect detection of raised fabrics based on improved EfficientNet** [13396-15]

13396 0W **Reconstructing the motion trajectory of an intelligent cart based on an upper unit** [13396-2]

13396 0X **Research on tree detection method based on UAV LIDAR point cloud** [13396-4]

13396 0Y **Small crop detection based on improved YOLOv8 algorithm for robotic pollination** [13396-9]

13396 0Z **Research on goat herd tracking in natural scenes based on improved DeepSORT** [13396-22]

13396 10 **Research and application based on the improved YOLO V7 target detection algorithm** [13396-10]