

PROCEEDINGS OF SPIE

***Third International Conference on  
Computer Vision and Pattern  
Analysis (ICCPA 2023)***

**Linlin Shen  
Guoqiang Zhong**  
*Editors*

**7–9 April 2023  
Hangzhou, China**

*Organized by*  
Shenzhen University (China)

*Sponsored by*  
AEIC Academic Exchange Information Centre (China)

*Published by*  
SPIE

**Volume 12754**

**Part One of Two Parts**

Proceedings of SPIE 0277-786X, V. 12754

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](http://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 *Third International Conference on Computer Vision and Pattern Analysis (ICCPA 2023)*, edited by Linlin Shen, Guoqiang Zhong, Proc. of SPIE 12754, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9781510667563

ISBN: 9781510667570 (electronic)

Published by

**SPIE**

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

Telephone +1 360 676 3290 (Pacific Time)

[SPIE.org](http://SPIE.org)

Copyright © 2023 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](http://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](http://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*

## Part One

---

### COMPUTER VISION AND IMAGE MODELING

---

- 12754 02 **Digital multi-scale visual planning model of spatial-geographical landscape pattern of smart parks [12754-61]**
- 12754 03 **Visual question answering model based on fusing global-local feature [12754-132]**
- 12754 04 **Object size detection based on image processing [12754-182]**
- 12754 05 **Image processing of the special sensor microwave/imager based on passive microwave remote sensing [12754-175]**
- 12754 06 **Classification of non-motorized vehicle based on micro-Doppler effect and image processing [12754-31]**
- 12754 07 **Dynamic visual SLAM system based on weighted static features [12754-7]**
- 12754 08 **Real-time object detection for UAVs images based on improved YOLOv5 [12754-125]**
- 12754 09 **Graptolite image classification based on feature transfer and mixup data enhancement [12754-159]**
- 12754 0A **A UAV detection framework based on dual stereo vision system for low-altitude urban background [12754-51]**
- 12754 0B **Research on image classification based on neural network VGG16 [12754-128]**
- 12754 0C **English handwriting image edge detection based on fuzzy logic [12754-134]**
- 12754 0D **An image classification method based on few-shot learning [12754-49]**
- 12754 0E **A lung cancer tumor image segmentation method of a CFC-MSPCNN based on PET/CT [12754-131]**
- 12754 0F **Fine-grained image recognition based on multi-branch and multi-scale learning [12754-90]**

- 12754 OG **Research on road extraction model of remote sensing image based on the fused convolutional module and attention mechanism** [12754-122]
- 12754 OH **A novel speckle-robust feature map and research on rotation invariance of feature maps in the SAR image** [12754-42]
- 12754 OI **Unsupervised aircraft detection in SAR images with image-level domain adaption from optical images** [12754-163]
- 12754 OJ **The role of echocardiography segmentation evaluation metrics in clinical diagnosis** [12754-120]
- 12754 OK **DAT-SegNet: a medical multi-organ segmentation network based on deformation attention and transformer** [12754-81]
- 12754 OL **Study on the classification method of gastrointestinal endoscope image based on CBAM and AlexNet** [12754-117]
- 12754 OM **Parallel architecture based on transformers and CNNs for renal cell carcinoma image segmentation** [12754-154]
- 12754 ON **Research on multi-view stereo based on depth estimation** [12754-115]
- 12754 OO **Machine vision-based measurement of air compressor crankshaft journal dimensions** [12754-114]
- 12754 OP **Analysis of machine vision-based defect detection for logistics conveyor belts** [12754-91]
- 12754 OQ **Design of a portable calibration device for medical endoscopy** [12754-36]
- 12754 OR **Lightweight face liveness detection method based on 32×24 thermal imaging images and ShuffleNet** [12754-47]
- 12754 OS **Assessment of depression level based on multi-scale 3D convolution and expression videos** [12754-48]
- 12754 OT **Design and implementation of video recording storage matrix based on domestic embedded SOC** [12754-127]
- 12754 OU **A remote sensing image change detection method incorporating variable convolution and attention mechanism** [12754-53]
- 12754 OV **Interactive image segmentation method based on asymmetric quadratic metric** [12754-41]
- 12754 OW **Transformer-based multi-scale gradient feature fusion for low-light image enhancement** [12754-84]
- 12754 OX **External attention augmented Swin-UNet network for semi-supervised cardiac image segmentation** [12754-104]
- 12754 OY **DBA: downsampling-based adversarial attack in medical image analysis** [12754-129]

- 12754 0Z **DCP-YOLOv7: dark channel prior image-dehazing YOLOv7 for vehicle detection in foggy scene** [12754-156]
- 12754 10 **Research on smoke removal of industrial scene image based on an improved dark channel prior** [12754-116]
- 12754 11 **A video vibration measurement core based on FPGA** [12754-86]

---

#### MOBILE WEB SERVICE AND INFORMATION DETECTION

---

- 12754 12 **The study of the intelligent voice man-machine interaction on the display console** [12754-171]
- 12754 13 **Automated inspection of member deformation in grid structures using three-dimensional laser scanning** [12754-17]
- 12754 14 **OpenPose-based model for infant cerebral palsy auxiliary detection** [12754-176]
- 12754 15 **A lightweight network for crack detection with lightweight feature extraction and multi-scale feature fusion** [12754-69]
- 12754 16 **Dynamic target detection based on YOLOv5** [12754-6]
- 12754 17 **Design and exploration of a virtual reality-based human-computer interaction system** [12754-183]
- 12754 18 **Research on trash detection based on instance segmentation** [12754-76]
- 12754 19 **The eligibility detection method of runway edge lights based on improved CenterNet** [12754-95]
- 12754 1A **Siamese tracker based on hybrid dilated convolution and global cross-correlation** [12754-24]
- 12754 1B **Research on global illumination exploration in dynamic scenes** [12754-168]
- 12754 1C **Dense pedestrian detection based on multiple link feature pyramid networks** [12754-164]
- 12754 1D **Facial expression transfer based on generative adversarial networks: a review** [12754-73]
- 12754 1E **Spectral fitting study of the CIE standard illuminant D65 based on monochromatic LED** [12754-63]
- 12754 1F **Complete imaging and distortion evaluation of spherical fruit surface based on concave spherical mirrors** [12754-82]
- 12754 1G **Semantic segmentation of point cloud based on graph neural network** [12754-15]

- 12754 1H **An adaptive inherent bias correction strategy embedded with spatio-temporal attention network for pedestrian trajectory prediction** [12754-28]
- 12754 1I **Pedestrian multi-object tracking based on YOLOv7 and BoT-SORT** [12754-68]
- 12754 1J **Road crack detection based on improved U-Net** [12754-92]
- 12754 1K **Spatial target detection method based on frame difference and target characteristics** [12754-158]
- 12754 1L **A parallel sea-land segmentation method based on improved UNet model** [12754-119]
- 12754 1M **Lightweight concrete crack segmentation method based on Deeplabv3+** [12754-94]
- 12754 1N **An efficient harbor ship detection method based on deep learning in high-resolution remote sensing images** [12754-162]
- 12754 1O **Design exploration of a tactile path defect detection vehicle** [12754-38]
- 12754 1P **Dynamic SLAM system based on object detection** [12754-26]
- 12754 1Q **Freeway vehicle detection on improved YOLOv7** [12754-21]
- 12754 1R **Automated depression diagnosis based on facial expression and deep dual-stream convolutional neural networks** [12754-32]
- 12754 1S **Extraction of soybean planting areas based on multi-temporal Sentinel-1/2 data** [12754-8]
- 12754 1T **A novel UAV target detection method based on the improved YOLOv4-tiny network** [12754-30]
- 12754 1U **EFFNet urban point cloud classification method** [12754-39]
- 12754 1V **Research on real-time detection model of chainsaw chain based on improved YOLOv5s** [12754-29]
- 12754 1W **Apple automatic classification method based on improved VGG11** [12754-16]
- 12754 1X **Collision detection of sensorless cooperative robots based on long short term memory network** [12754-160]
- 12754 1Y **Hierarchical feature fusion network with receptive field block for fabric defect detection** [12754-43]
- 12754 1Z **Cigarette multi-view labelling forecasting in the new retail model** [12754-143]
- 12754 20 **Nanodet-Ghost: a lightweight network for quality detection of wheat kernels appearance** [12754-186]

12754 21 **Detection of farm harvestable apples based on improved YOLOv7** [12754-72]

## **Part Two**

### **COMPUTER ALGORITHM AND DATA ANALYSIS**

---

12754 22 **Research on user's facial expression analysis when watching TV** [12754-78]

12754 23 **Double-frame rate-based cheat detection** [12754-88]

12754 24 **Adaptive multi-modal decision fusion for RGB-T tracking** [12754-3]

12754 25 **Automatic crack detection and calculation based on deep learning and digital image processing** [12754-184]

12754 26 **Research on image stitching algorithm based on improved SURF features** [12754-181]

12754 27 **Research on efficient image semantic segmentation algorithm for glass transparent containers based on improved DeepLabV3+** [12754-172]

12754 28 **A dynamic SLAM algorithm for fusion of semantic and depth information** [12754-83]

12754 29 **Multi-source image matching using improved ORB algorithm** [12754-27]

12754 2A **PASSD: an improved SSD object detection algorithm based on feature fusion with feedback-based loss for small object detection** [12754-106]

12754 2B **Research on stereo matching algorithm of underwater binocular detection** [12754-136]

12754 2C **An efficient algorithm for detecting shortage of soybean seedlings using RGB images captured by UAVs** [12754-71]

12754 2D **A lightweight face detection algorithm for wearing masks** [12754-98]

12754 2E **Research on leading scan technology and its application in NDT of nuclear power** [12754-118]

12754 2F **Object detection algorithm based on mixed dilated convolution pyramid** [12754-169]

12754 2G **Calculation method of road distress area based on monocular vision** [12754-113]

12754 2H **Face anti-spoofing based on ASFF and LBP self-supervision branch** [12754-167]

12754 2I **Improved-GCN-based collaborative filtering recommendation model** [12754-35]

12754 2J **Segmentation of sandstone particles based on improved watershed algorithm** [12754-85]

- 12754 2K **A novel anti-occlusion and scale adaptation re-tracking method with correlation filtering tracker algorithm** [12754-110]
- 12754 2L **Improved clothing segmentation algorithm based on Deeplabv3 + network** [12754-126]
- 12754 2M **An improved Apriori algorithm based on influence weight** [12754-60]
- 12754 2N **Non-ellipsoidal extended object tracking using PMBM filter** [12754-23]
- 12754 2O **Research on collaborative filtering recommendation algorithm with improved K-means clustering** [12754-138]
- 12754 2P **Multi-population particle swarm optimization concerning the worst information** [12754-109]
- 12754 2Q **Research on daily behavior recognition method based on computer vision** [12754-50]
- 12754 2R **Dynamic hand gesture classification based on attention mechanism and transformer** [12754-25]
- 12754 2S **Analysis of vehicle management based on an improved K-prototype** [12754-111]
- 12754 2T **A binocular stereo vision packaging dimension measurement algorithm based on improved SURF** [12754-74]
- 12754 2U **Noise reduction method for seismic data based on subspace projection and fast Fourier convolution** [12754-59]
- 12754 2V **Research on arbitrary style transfer algorithm based on attention mechanism** [12754-52]
- 12754 2W **Study on extraction of key features of sugar orange phenotype** [12754-139]
- 12754 2X **A study on multi-strategy improvement algorithm for chimpanzee optimization by fusing differential variation and random variation** [12754-166]

---

#### DEEP LEARNING AND ARTIFICIAL INTELLIGENCE RECOGNITION

- 12754 2Y **Road detection and three-dimensional visualization management using deep learning and BIM integration** [12754-5]
- 12754 2Z **Traffic sign detection based on deep learning** [12754-177]
- 12754 30 **Face gesture recognition module optimization of intelligent security system based on Raspberry Pi** [12754-9]
- 12754 31 **Application of a semantic segmentation method based on deep learning in unity scene construction** [12754-2]



- 12754 32 **Underwater acoustic target recognition based on convolutional neural network and multi-feature fusion** [12754-157]
- 12754 33 **Design and implementation of an optimization human pose estimation scheme** [12754-4]
- 12754 35 **Research on image target recognition of defects in steel structure welded joints** [12754-185]
- 12754 36 **Research on intelligent discrimination technology of conductor insulator damage based on UAV images** [12754-151]
- 12754 37 **Abnormal recognition of the ETC gantry license plate data based on an improved DBSCAN** [12754-20]
- 12754 38 **Rotated vehicle object detection and tracking on drone-captured scenarios** [12754-112]
- 12754 39 **Research on flame detection of fire-fighting robot** [12754-14]
- 12754 3A **Adaptive and learnable label for face anti-spoofing** [12754-170]
- 12754 3B **Handwritten mathematical expression recognition based on ResNet and transformer** [12754-19]
- 12754 3C **A driving style recognition method based on SAX and bitmap** [12754-58]
- 12754 3D **Multi-angle truck model recognition based on detail awareness** [12754-18]
- 12754 3E **EEG emotion recognition technology for children with autism** [12754-130]
- 12754 3F **Aesthetic evaluation of handwritten Chinese characters based on deep learning** [12754-44]
- 12754 3G **Storage optimization method based on region of interest in part pose recognition** [12754-33]
- 12754 3H **Research on image recognition technology based on CNN** [12754-97]
- 12754 3I **Deep learning-based defect detection for printed circuit boards** [12754-70]
- 12754 3J **Traffic police command gesture recognition technology based on machine vision and two-stream spatio-temporal attention graph convolutional network** [12754-12]
- 12754 3K **Facial expression optimal class separability for expression recognition** [12754-57]
- 12754 3L **Design and implementation of intelligent soccer-picking robot** [12754-161]
- 12754 3M **Research on vehicle logo positioning method based on deep learning** [12754-96]
- 12754 3N **iGuard: an intelligent sitting posture monitoring system with pressure sensors** [12754-137]

- 12754 3O **Traffic sign recognition based on improved VGG16 algorithm** [12754-11]
- 12754 3P **Gait recognition based on local relationship learning** [12754-75]
- 12754 3Q **New progress in brain-like collaborative control based on brain inspiration** [12754-152]
- 12754 3R **Method of feature extraction and recognition for surface defect of sugar orange** [12754-146]
- 12754 3S **Co-teaching module for multiple instance learning-based weakly supervised object detection network** [12754-123]
- 12754 3T **Combining autoencoder and category-based low-rank domain adaptation method for multi-site ASD identification** [12754-135]
- 12754 3U **Research on detection and recognition system of oil and gas pipeline weld seam based on machine vision technology** [12754-149]
- 12754 3V **Study on grasping pose detection based on lightweight artificial neural network** [12754-54]
- 12754 3W **An overview of Chinese fine-grained sentiment analysis methods based on multi-kernel plant intelligence model** [12754-107]
- 12754 3X **Research on hand sketch recognition method based on Sketch-AlexNet** [12754-101]
- 12754 3Y **Research on vehicle speech recognition algorithm based on multidimensional data** [12754-13]