

PROCEEDINGS OF SPIE

Fourth International Conference on Computer Vision and Data Mining (ICCVDM 2023)

Chunyi Chen
Xin Zhang
Editors

20–22 October 2023
Changchun, China

Organized by
Changchun University of Science and Technology (China)

Sponsored by
AEIC Academic Exchange Information Centre (China)

Published by
SPIE

Volume 13063

Proceedings of SPIE 0277-786X, V. 13063

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 *Fourth International Conference on Computer Vision and Data Mining (ICCVDM 2023)*, edited by Chunyi Chen, Xin Zhang, Proc. of SPIE 13063, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9781510674448

ISBN: 9781510674455 (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**

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 Committee*

PATTERN RECOGNITION AND VISUAL FEATURE RESEARCH

- 13063 02 **Fruit image classification based on ResNet50 and transfer learning** [13063-78]
- 13063 03 **Image recognition model building in the context of municipal waste classification** [13063-107]
- 13063 04 **Equilibrium and integrity correction class activation map in weakly supervised semantic segmentation** [13063-108]
- 13063 05 **Intelligent tunnel fire emergency linkage measures based on image recognition** [13063-30]
- 13063 06 **Wildlife detection and identification based on the improved YOLOv7** [13063-113]
- 13063 07 **A method for the super-resolution reconstruction of remote sensing images based on multi-scale feature fusion and an attention mechanism** [13063-124]
- 13063 08 **Robot operating system for live work in substation based on machine vision** [13063-16]
- 13063 09 **Remote sensing image small target and moving target change detection method** [13063-83]
- 13063 0A **People recognition with foot pressure based on neural network** [13063-40]
- 13063 0B **MVS-STRNet: cross-view space target recognition from multi-view stereo** [13063-70]
- 13063 0C **Video visibility prediction method for foggy weather based on spatio-temporal variation features** [13063-45]
- 13063 0D **Comparison and improvement of some attention mechanisms in the video field** [13063-105]
- 13063 0E **Enhancing video quality assessment through brightness information** [13063-130]
- 13063 0F **ERBNet: an automatic polyp image segmentation method to enhance receptive field and boundary** [13063-14]
- 13063 0G **Exposition and implementation of linear equation system reconstruction method in electromagnetic wave CT imaging** [13063-1]
- 13063 0H **Deep learning-based 3D face reconstruction method for video stream** [13063-118]

- 13063 0I **Video anomaly detection via improved future frame prediction** [13063-22]
- 13063 0J **Image colorization based on transformer with sparse attention** [13063-92]
- 13063 0K **SProtoNet: self-supervised ProtoNet for plant leaf disease few-shot classification** [13063-23]
- 13063 0L **Optimal use of attention mechanisms: comparative study in U-Net for image segmentation tasks** [13063-99]
- 13063 0M **GDformer: a lightweight decoder for efficient semantic segmentation of remote sensing urban scene imagery** [13063-87]
- 13063 0N **In-vehicle fatigue driving warning device based on facial recognition fatigue detection system** [13063-4]
- 13063 0O **Image alignment-based patch distribution framework for anomaly detection** [13063-100]
- 13063 0P **Substation anomaly detection method based on attention mechanism and feature fusion** [13063-120]
- 13063 0Q **Automatic speech recognition based on attention-enhanced blockformer** [13063-86]
- 13063 0R **Research on the method of accurate recognition of target object pose by intelligent robot using CDPN in visual guidance** [13063-93]
- 13063 0S **Automatic speech recognition based on conformer with convolution enhanced channel attention block** [13063-112]
- 13063 0T **Research on the application of image-based analysis in interaction design** [13063-50]
- 13063 0U **Embedded surface crack image classification based on multi-model voting** [13063-41]
- 13063 0V **Small target detection based on attention mechanism feature fusion** [13063-54]
- 13063 0W **Image fusion with inverted residual densely connected directional differentiation and attention mechanisms** [13063-117]
- 13063 0X **An improved YOLOv8s model based on sphere convolution for object detection in fisheye images** [13063-37]
- 13063 0Y **Video semantic segmentation based on spatio-temporally polarized attention** [13063-74]
- 13063 0Z **A study on the method of recognizing star maps in the all-sky region** [13063-106]
- 13063 10 **A continual learning-based approach for multi-source image enhancement and reconstruction over different categorical scenes** [13063-101]
- 13063 11 **Application of transformer on the multitask of identification of accent and gender toward an audio** [13063-49]

13063 12 **Biomedical named entity recognition based on reinforcement learning label correction**
[13063-58]

INTELLIGENT INFORMATION RETRIEVAL AND DECISION ANALYSIS

13063 13 **Research on new media recommendation algorithm based on comment text** [13063-29]

13063 14 **Shadow recovery based on panorama complex environment** [13063-98]

13063 15 **Brain tumor segmentation by combining MRI multimodal information and improved Unet network** [13063-65]

13063 16 **Efficient distribution solution for VBR encoded video** [13063-131]

13063 17 **Landmark extraction for hippocampal surfaces** [13063-2]

13063 18 **A visually secure image encryption scheme based on parallel compressive sensing**
[13063-133]

13063 19 **Quantization parameter decision scheme of HEVC** [13063-43]

13063 1A **Design and implementation scheme of evaluation board based on STM32 for SFP optical module sensitivity test** [13063-60]

13063 1B **High permeability flowing channel prediction upon production data mining** [13063-123]

13063 1C **FF-DETR: feature-fusion transformers for end-to-end object detection** [13063-56]

13063 1D **Enhanced correlation filter tracking: anti-occlusion and re-detection** [13063-57]

13063 1E **Research on color quality classification algorithms in computer art education** [13063-71]

13063 1F **Foreign object detection system in catenary networks based on novel cascade convolutional neural networks** [13063-42]

13063 1G **Privacy preserving machine learning in energy services: a survey** [13063-114]

13063 1H **Efficiency study of VGG networks in autonomous driving tasks** [13063-66]

13063 1I **Research on position and orientation measuring method of auto-landing based on machine vision** [13063-67]

13063 1J **Leveraging behavioral data mining for cost reduction and efficiency enhancement in manufacturing** [13063-69]

- 13063 1K **Secure transmission of sensing data based on Arnold transform and steganography technique** [13063-125]
- 13063 1L **Integrating the Inception V3 model in an end-to-end driving scenario** [13063-32]
- 13063 1M **A denoising model for MC rendered images based on fusion kernel prediction and generation of adversarial networks** [13063-127]
- 13063 1N **VeinCnnformer: convolutional neural network based transformer for vein recognition** [13063-126]
- 13063 1O **FedCOA: an efficient cluster optimization for federated learning** [13063-7]
- 13063 1P **Research on the evaluation of signpost design schemes using physiological measurement methods** [13063-59]
- 13063 1Q **Research on image features extraction and classification using transformer model** [13063-128]
- 13063 1R **Point-voxel feature set abstraction with hybrid cross feature fusion module for 3D object detection** [13063-72]
- 13063 1S **Entity relation joint extraction model combining pointer network and attention mechanism based on relative position embedding** [13063-104]
- 13063 1T **Research on extractive text abstract generation method for graph model based on TextRank** [13063-48]
- 13063 1U **Research on course teaching evaluation and academic early warning based on data mining** [13063-82]
- 13063 1V **OICR-APIM: weakly supervised object detection with integrated global attention and pseudo-instance mining** [13063-94]
- 13063 1W **We-Kcr: a new protein crotonylation site prediction method based on sequence feature and neural network** [13063-36]
- 13063 1X **Satellite pose measurement using an improving SIFT algorithm** [13063-24]
- 13063 1Y **Multivariate time series forecasting based on fractional Fourier transforms** [13063-96]
- 13063 1Z **Lightweight correlation filter based object tracking network** [13063-8]
- 13063 20 **Time-domain speech separation based on local parallel recursive networks and global fast attention** [13063-102]

INTELLIGENT TARGET DETECTION AND ALGORITHM FUSION

- 13063 21 **An insulator detection method based on optimized Yolov5s** [13063-27]

- 13063 22 **A smart lighting system based on digital twin technology in smart home** [13063-51]
- 13063 23 **Research on improved Yolov5 license plate location technology** [13063-95]
- 13063 24 **Style transfer based adversarial patches for object detection model attacks** [13063-5]
- 13063 25 **Efficient multi-instance hashing for large-scale image retrieval** [13063-55]
- 13063 26 **Fast algorithm for intra prediction mode of 360 degree panoramic video based on VVC**
[13063-111]
- 13063 27 **Human fall detection algorithm based on random forest and MPU6050** [13063-79]
- 13063 28 **Unmanned aerial vehicles group target detection algorithm based on group structure**
[13063-62]
- 13063 29 **A road crack detection method based on improved U-Net** [13063-9]
- 13063 2A **A lightweight vehicle-pedestrian detection algorithm based on YOLOv5s** [13063-26]
- 13063 2B **Image detection based on the grey wolf algorithm optimization and its application in campus security** [13063-15]
- 13063 2C **Fruit detection and recognition with deep learning** [13063-53]
- 13063 2D **Research on improving the detection method of welding defects in drum welding of CoT-YOLOv7** [13063-89]
- 13063 2E **Two-stage small target detection algorithm based on feature fusion** [13063-28]
- 13063 2F **Near-infrared spectroscopy based on transfer learning, detection of tobacco chemical composition** [13063-80]
- 13063 2G **A vehicle detection method based on Taylor expansion model pruning algorithm** [13063-39]
- 13063 2H **Detection of dim small targets in strong backgrounds based on local feature maximum difference** [13063-75]
- 13063 2I **Polarization fusion algorithm based on NSCT decomposition and improved dual-channel PCNN model** [13063-77]
- 13063 2J **Enhancing accuracy and efficiency of small object detection in remote sensing an optimized YOLOv5 approach** [13063-46]
- 13063 2K **Innovative application of virtual reality technology in digital display of intangible cultural heritage** [13063-19]

- 13063 2L **Research on 3D object detection for point cloud** [13063-64]
- 13063 2M **Image watermarking algorithm based on localized FAPHFs magnitude** [13063-21]
- 13063 2N **Research on a wheel hub quality judgment processing algorithm** [13063-17]
- 13063 2O **A mutual nearest neighbor search algorithm based on dynamic neighborhood** [13063-115]
- 13063 2P **Robust iris localization algorithm based on improved YOLOv4 and self-organized particle swarm optimization** [13063-129]
- 13063 2Q **Driver anomalous driving behavior detection based on supervised contrastive learning** [13063-13]
- 13063 2R **The attribute classification model in intelligent question answering system based on domain knowledge graph** [13063-91]
- 13063 2S **Convolution pyramid attention: an efficient channel attention mechanism** [13063-25]
- 13063 2T **Research on small target detection algorithm for UAV inspection scene based on SPD-conv** [13063-81]
- 13063 2U **Polarization image fusion algorithm based on gradient residual generation adversarial network** [13063-52]
- 13063 2V **The graph attention network in relational reasoning based on knowledge graph** [13063-90]
- 13063 2W **Improving vehicle classification model algorithm based on ImageNet pretraining** [13063-85]
- 13063 2X **A study on multidimensional evaluation of teaching using loose-leaf textbooks based on improved K-modes algorithm** [13063-97]
- 13063 2Y **Harnessing artificial intelligence for ophthalmic disease diagnosis: a comparative study of CNNs and Swin transformer models** [13063-84]
- 13063 2Z **Improved sparrow search particle filtering target tracking algorithm** [13063-68]