Eighth International Conference on Video and Image Processing (ICVIP 2024)

Xuefeng Liang Editor

13–15 December 2024 Kuala Lumpur, Malaysia

Organized by Xidian University (China) Singapore Institute of Electronics (Singapore)

Published by SPIE

Volume 13558

Proceedings of SPIE 0277-786X, V. 13558

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 Eighth International Conference on Video and Image Processing (ICVIP 2024), edited by Xuefeng Liang, Proc. of SPIE 13558, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

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

ISBN: 9781510689237 ISBN: 9781510689244 (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.



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

-

v Conference Committee

TARGET DETECTION AND APPLICATION TECHNOLOGY

neurological and cardiac signals [13558-13]

13558 02	Shape-guided unsupervised 3D defect detection [13558-15]
13558 03	Depth-aware weather-enhanced 3D object detection for roadside infrastructure [13558-9]
13558 04	Research on lightweight algorithm for cargo packaging unsealing detection based on improved YOLOv5 [13558-17]
13558 05	WTConv-Adapter: wavelet transform convolution adapter for parameter efficient transfer learning of object detection ConvNets [13558-3]
13558 06	Learning discontinuity in orientation with cosine metric in real-oriented object detection [13558-18]
	FEATURE DETECTION AND FUSION
13558 07	A multi-location enhanced multi-scale feature fusion skin lesion segmentation network based on U-shaped network [13558-19]
13558 08	Railway clearance intrusion detection using feature fusion enhancement neural network [13558-1]
13558 09	Progressive self-supervised spatio-temporal feature learning based on video sequence saliency [13558-2]
	IMAGE RECOGNITION AND ALGORITHM
13558 0A	VIVAR: learning view-invariant embedding for video action recognition [13558-11]
13558 OB	A student behavior recognition algorithm based on improved MobileNetV2 [13558-4]
13558 OC	Survey on emotion recognition systems based on facial expressions combined with

iii

COMPUTER IMAGING AND IMAGE MODEL

- 13558 0D Does display size affect short-term memory tasks? [13558-21]
- 13558 OE Dual photography using hierarchical orthogonal codes [13558-22]
- 13558 OF SPGMS: an annotated benchmark for grain-matrix segmentation in sandstone photomicrographs [13558-6]
- 13558 0G Research on the application of VR in the learning of different algorithm models and the construction of spatial interaction models [13558-8]

DIGITAL IMAGE ANALYSIS AND PROCESSING METHOD

- 13558 0H An attention-guided image super-resolution for scene text [13558-7]
- 13558 01 S-DIHE: secure deduplication of images based on homomorphic encryption [13558-20]
- 13558 0J Point-plane better than point-point for embedding-based USAD [13558-5]
- 13558 OK Deep learning framework for analysing land use and land cover classification and change detection [13558-24]