

PROCEEDINGS OF SPIE

Sixteenth International Conference on Graphics and Image Processing (ICGIP 2024)

Liang Xiao
Editor

8–10 November 2024
Nanjing, China

Organized by
Nanjing University of Science and Technology (China)
Jiangsu Computer Society (China)
Jiangsu Association of Artificial Intelligence (China)

Sponsored by
Nanjing University of Science and Technology (China)

Cosponsored by
Soochow University (China)
Suzhou University of Science and Technology (China)
Nanjing University of Posts and Telecommunications (China)
Nanjing University of Information Science and Technology (China)

Published by
SPIE

Volume 13539

Proceedings of SPIE 0277-786X, V. 13539

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 *Sixteenth International Conference on Graphics and Image Processing (ICGIP 2024)*, edited by Liang Xiao, Proc. of SPIE 13539, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9781510688773

ISBN: 9781510688780 (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

ix	Conference Committee
xiii	Introduction

OBJECT DETECTION AND ALGORITHMS

13539 02	A novel graph attention network for underwater object detection [13539-80]
13539 03	OGSOD-2.0: A challenging multimodal benchmark for optical-SAR object detection [13539-36]
13539 04	Saliency guided RT-DETR for object detection [13539-95]
13539 05	Visual SLAM algorithm based on lightweight object detection network in dynamic environments [13539-60]
13539 06	Deformable midline shift network: automatic brain-midline shift detection using dynamic snake convolution [13539-27]
13539 07	Enhanced prototypical contrastive learning for carbon fiber defect detection [13539-39]
13539 08	Few-shot object detection via query feature calibration [13539-99]
13539 09	Remote sensing few-shot object detection using adaptive feature modification and Gaussian dynamic dilated convolution [13539-82]
13539 0A	Research on a portable automatic aircraft surface damage detection technology [13539-81]
13539 0B	Mask-patchcore: a robust anomaly detection model focusing on interested region [13539-74]
13539 0C	Monocular 3D object detection via object edge information [13539-34]
13539 0D	A driver fatigue detection method based on improved Kalman filter algorithm under infrared conditions [13539-114]
13539 0E	Camouflaged object detection via local features and selective fusion [13539-35]

IMAGE SEGMENTATION

- 13539 OF **Double-click: dual-click iteration and dual-branch networks for interactive segmentation** [13539-26]
- 13539 OG **A novel multiscale 3D segmentation network based on vascular skeleton assistance** [13539-47]
- 13539 OH **A stroke lesion segmentation method based on multi-scale and cascade attention mechanisms** [13539-44]
- 13539 OI **MedSegKAN: a superior medical image segmentation method based on the improved KAN structure** [13539-43]
- 13539 OJ **Global-local transformer for aerial image semantic segmentation** [13539-40]
- 13539 OK **Deformable coordinate kernel attention-based network for medical image segmentation** [13539-68]
- 13539 OL **Patch-free tooth segmentation from cone beam computed tomography dataset** [13539-12]
- 13539 OM **Fast, trusted semantic segmentation of LiDAR point clouds with evidential fusion** [13539-93]
- 13539 ON **DPDT: semi-supervised choroidal vessel segmentation via deformation perception and dynamic teacher selection** [13539-9]
- 13539 OO **Multimodal knowledge distillation-guided artery-vein segmentation network for OCT/OCTA images** [13539-8]

IMAGE CLASSIFICATION

- 13539 OP **Discriminative maximum mean and covariance discrepancy based domain adaptation for cross-domain image classification** [13539-58]
- 13539 OQ **Distance aware text optimizer for multimodal prompt learning** [13539-65]
- 13539 OR **Hybrid spectral-spatial ResNet and transformer for hyperspectral image classification** [13539-108]
- 13539 OS **Abnormal uterine classification based on an improved YOLOv5 framework from ultrasound images** [13539-117]
- 13539 OT **Spatial vs. spectral resolution: which is important for lychee and longan classification from hyperspectral remote sensing image** [13539-10]
- 13539 OU **An image classification algorithm based on the topological features** [13539-79]

INTELLIGENT RECOGNITION AND ALGORITHMS

- 13539 0V **MV-ResFPN: a residual network deeply combining multi-view and multiscale Fusion for SAR target recognition** [13539-56]
- 13539 0W **Prototype transferred diverse features for day-night cross-domain facial expression recognition** [13539-118]
- 13539 0X **Research on face recognition algorithm based on ELPP and fuzzy k-nearest neighbor algorithm** [13539-42]
- 13539 0Y **A UAV object tracking and counting method based on change identity and boundary awareness** [13539-111]
- 13539 0Z **Research on writer identification based on attention mechanism** [13539-86]

IMAGE MODELING AND IMAGE RECONSTRUCTION

- 13539 10 **Asymmetric lens distortion compensation method based on linear reconstruction of feature points** [13539-38]
- 13539 11 **Self-supervised multiview reconstruction of eyeglasses frames with thin structures** [13539-3]
- 13539 12 **Intelligent depiction of light and shadow** [13539-88]
- 13539 13 **Point cloud reconstruction and deformation based on point primitives** [13539-66]
- 13539 14 **Double partial said-ball curves: a new type of generalized ball curves** [13539-73]
- 13539 15 **Single image super-resolution reconstruction based on multiscale attention mechanism** [13539-77]
- 13539 16 **CALFNet: a light field reconstruction method based on channel attention mechanism** [13539-57]
- 13539 17 **A nonconvex TVcapped-L1 model for image restoration with impulse noise** [13539-13]

MULTIMODAL PERCEPTION AND FEATURE FUSION TECHNIQUES IN IMAGE PROCESSING

- 13539 18 **Foreground prediction for image composition with local and global feature fusion** [13539-96]
- 13539 19 **Improved feature extraction and matching algorithm for point-line fusion** [13539-49]

- 13539 1A **TSAM: Multilevel condition-enhanced image inpainting via progressive feature fusion and decoding mechanism** [13539-51]
- 13539 1B **Dense video event semantic annotation based on multimodal features** [13539-41]
- 13539 1C **Fuzzy-AttnNet: a deep learning framework for image retrieval using fuzzy feature fusion** [13539-21]
- 13539 1D **Point cloud completion via trigonometric encoding and self-attention-based feature fusion** [13539-7]
- 13539 1E **Hyperspectral change detection based on 3D convolution autoencoder and feature difference constraint** [13539-16]

COMPUTER AIDED PAINTING AND MULTIMEDIA APPLICATIONS

- 13539 1F **Optics guided thermal image super-resolution via alignment fusion network** [13539-94]
- 13539 1G **Efficient random methods for low-rank matrix estimation** [13539-50]
- 13539 1H **Line art extraction of Chinese building structure based on ControlNet model** [13539-72]
- 13539 1I **Refining boundary sample space for differentiable rendering** [13539-89]
- 13539 1J **BoxNeRF: simplified object selection and editing in NeRF scenes** [13539-76]
- 13539 1K **BEOUNet: video frame interpolation via bidirectional encoding and optimized U-Net** [13539-59]
- 13539 1L **A method for the visual representation of implicit surface animation based on temporal interval inversion** [13539-98]
- 13539 1M **Research on the automatic calculation of surface area integrals** [13539-102]

IMAGE PROCESSING TECHNOLOGY BASED ON DEEP LEARNING AND MULTIMODAL PERCEPTION

- 13539 1N **An adaptive geodesic voting method for curvilinear tree structure extraction** [13539-46]
- 13539 1O **Underwater image enhancement via multiple attention mechanisms and adversarial learning** [13539-85]
- 13539 1P **Reversible information hiding algorithm based on classification of SMVQ indices** [13539-20]
- 13539 1Q **An image compression algorithm based on data embedding** [13539-37]

- 13539 1R **WGID: wavelet-guided iterative detail enhancement diffusion models for zero-shot image super-resolution** [13539-53]
- 13539 1S **High-resolution forward-looking MIMO-SAR imaging approach for automotive mmwave radar** [13539-62]
- 13539 1T **A wavelet-based diffusion model for low-light image enhancement in raw space** [13539-33]
- 13539 1U **A novel gray code assisted phase measuring profilometry for high-speed objects** [13539-2]

AI BASED IMAGE GENERATION AND RESTORATION TECHNOLOGY

- 13539 1V **An improved patch-based inpainting method for ancient mural images** [13539-90]
- 13539 1W **Enhanced reversible data hiding in encrypted images by pixel shifting** [13539-4]
- 13539 1X **Evaluating pixel pair for image matting using hybrid feature distortion** [13539-101]
- 13539 1Y **Calligraphic guideline generation via contour-guided brush fitting** [13539-87]
- 13539 1Z **Texture generation with consistent lighting and multiview content via AIGC** [13539-1]
- 13539 20 **Latent age attribute modulation guided continuous aging facial image generation** [13539-45]
- 13539 21 **Human novel pose generation with pose optimization and importance sampling based on NeRF** [13539-55]
- 13539 22 **SAR image generation method via improved CWGAN-GP** [13539-17]
- 13539 23 **Real-time generation of digital Chinese calligraphy based on geometric continuity** [13539-15]

COMPUTER VISION AND IMAGE ANALYSIS

- 13539 24 **Dynamic visual SLAM based on improved ORB-SLAM3 tracking thread** [13539-54]
- 13539 25 **Hierarchical matching for visible and infrared images of power equipment based on multiscale local normalized filtering** [13539-67]
- 13539 26 **A fusion label for keypoint matching on multi-spectral images** [13539-83]

- 13539 27 **MoPoSD: Monocular 3D object detector with potential for specialized exploration in depth** [13539-75]
- 13539 28 **LCRNet: Unsupervised non-rigid point cloud registration network based on local correspondence relationships** [13539-5]
- 13539 29 **Deformable object manipulation based on double Q-map PPO** [13539-30]
- 13539 2A **Refinement for image rain removal based on diffusion model** [13539-103]
- 13539 2B **A study of adaptive sampling-based primitive fidelity methods** [13539-61]
- 13539 2C **Deep reasoning of large models based on knowledge graph** [13539-70]

DIGITAL IMAGE PROCESSING AND APPLICATIONS

- 13539 2D **Based on the improved YOLOv8n pose estimation algorithm** [13539-14]
- 13539 2E **Partial multilabel learning via dual subspace collaboration** [13539-31]
- 13539 2F **Transcriptional signatures of macroscopic cortical structure alternations in major depressive disorder** [13539-69]
- 13539 2G **Customized virtual liver tumor resection surgery simulation** [13539-32]
- 13539 2H **Exploring multimodal transformable imagery: a case study of human-AI co-creation in ICH** [13539-18]
- 13539 2I **Augmentative fusion network for robust RGBT tracking** [13539-52]
- 13539 2J **DKDTrack: dual-granularity knowledge distillation for RGBT tracking** [13539-48]
- 13539 2K **Development of an entry-exit decision algorithm for bus passenger counting using cameras** [13539-113]
- 13539 2L **Hyperspectral image methane detection based on DETR** [13539-91]
- 13539 2M **Pose estimation for non-cooperative satellites based on a three-dimensional bounding box detection network** [13539-78]
- 13539 2N **Multi-attribute fusion generation via text-to-image customization** [13539-100]
- 13539 2O **Density peak clustering algorithm based on fusing k-nearest neighbors and inverse neighbors** [13539-71]