PROCEEDINGS OF SPIE

Fourteenth International Conference on Digital Image Processing (ICDIP 2022)

Xudong Jiang
Wenbing Tao
Deze Zeng
Yi Xie
Editors

20–23 May 2022
Wuhan, China

Organized by
China University of Geosciences (China)

Sponsored by
Gifu University (Japan)
University of Poitiers (France)
University of New Hampshire (USA)

Published by
SPIE

Volume 12342

Part One of Two Parts

Proceedings of SPIE 0277-786X, V. 12342

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 *Fourteenth International Conference on Digital Image Processing (ICDIP 2022)*, edited by Xudong Jiang, Wenbing Tao, Deze Zeng, Yi Xie, Proc. of SPIE 12342, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

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

ISBN: 9781510657564
ISBN: 9781510657571 (electronic)

Published by
SPIE
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone +1 360 676 3290 (Pacific Time)
SPIE.org
Copyright © 2022 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

xiii *Conference Committee*

Part One

SESSION 1 FEATURE EXTRACTION AND FUSION

12342 02 **RSFS: A soft biometrics-based relative support features set for person verification** [12342-136]

12342 03 **LDAM: line descriptors augmented by attention mechanism** [12342-72]

12342 04 **Stochastic recursive gradient descent optimization-based on foreground features of Fisher vector** [12342-152]

12342 05 **An evaluation method of aggregate morphological characteristics based on two-dimensional digital image technique** [12342-148]

12342 06 **Semantic segmentation of road scene based on multi-scale feature extraction and deep supervision** [12342-158]

12342 07 **Dynamic spectral–spatial multiscale feature extraction network for hyperspectral image classification** [12342-27]

SESSION 2 TARGET DETECTION

12342 08 **An improved approach for two-stage detection model** [12342-20]

12342 09 **Multi-task model for human pose estimation and person detection** [12342-139]

12342 0A **Ship target detection method based on improved CenterNet in synthetic aperture radar images** [12342-106]

12342 0B **Lane detection with position embedding** [12342-102]

12342 0C **SAR target image edge detection based on CNN** [12342-43]

12342 0D **Ship detection in optical remote sensing images based on saliency and rotation-invariant feature** [12342-98]

12342 0E **Dock detection method in remote sensing images based on improved YOLOv4** [12342-17]

12342 0F **A defect detection method for plastic gears based on deep learning and machine vision** [12342-83]

12342 0G **Face tampering detection based on spatiotemporal attention residual network** [12342-153]

12342 0H **Ship target detection based on multitask learning** [12342-129]

12342 0I **Detecting of plant protection spraying by excited luminescence** [12342-173]

12342 0J **Deformation convolution and self-attention for fabric defect detection** [12342-154]

12342 0K **Using Sentinel-2 imagery for detecting oil spills via spatial roughness of mixed normalized difference index** [12342-155]

12342 0L **Breast cancer detection from ultrasound images using attention U-nets model** [12342-49]

12342 0M **RF2Net: salient object detection using level set loss and reverse attention fusion feed network** [12342-36]

12342 0N **Detecting roads from high-resolution aerial images: a position iteration algorithm for linear target detection** [12342-161]

SESSION 3 PATTERN RECOGNITION

12342 0O **Pre-rotation only at inference-time: a way to rotation invariance** [12342-119]

12342 0P **SlowFast with DropBlock and smooth samples loss for student action recognition** [12342-109]

12342 0Q **Spatio-temporal dual-attention network for view-invariant human action recognition** [12342-41]

12342 0R **Lightweight graph convolutional network with fusion data for skeleton based action recognition** [12342-65]

12342 0S **Masked facial region recognition using human pose estimation and broad learning system** [12342-14]

12342 0T **Fine-grained birds recognition based on lightweight bilinear CNN with Additive Margin Softmax** [12342-74]

12342 0U **A sample diversity and identity consistency based cross-modality model for visible-infrared person re-identification** [12342-103]

12342 0V **Application of artificial neural networks in recognizing carrier based on the color of raspberry powders obtained in the spray-drying process** [12342-172]

12342 0W **DIRA: disjoint-identity resolution adaptation for low-resolution face recognition** [12342-77]

12342 0X **Deep learning techniques for image recognition of counterfeited luxury handbags materials** [12342-156]

12342 0Y **A joint feature aggregation method for robust masked face recognition** [12342-51]

12342 0Z **Identification and localisation of multiple weeds in grassland for removal operation** [12342-85]

12342 10 **Rotated target recognition of sonar images via convolutional neural networks with rotated inputs** [12342-142]

SESSION 4 3D MODEL AND IMAGE RECONSTRUCTION

12342 11 **FEAFA+: an extended well-annotated dataset for facial expression analysis and 3D facial animation** [12342-48]

12342 12 **Grad-CAM based visualization of 3D CNNs in classifying fMRI** [12342-62]

12342 13 **Single image 3D scene reconstruction based on ShapeNet models** [12342-168]

12342 14 **ARO-DeepSFM: deep structure-from-motion with alternating recursive optimization** [12342-105]

12342 15 **Autism spectrum disorder analysis by using a 3D-ResNet-based approach** [12342-97]

12342 16 **3D face alignment and face reconstruction based on image sequence** [12342-138]

12342 17 **PointIt3D: a benchmark dataset and baseline for pointed object detection task** [12342-170]

12342 18 **Enabling deep reinforcement learning autonomous driving by 3D-LiDAR point clouds** [12342-108]

12342 19 **Reference-driven undersampled MRI reconstruction using automated stopping deep image prior** [12342-86]

12342 1A **Identifying Alzheimer's disease from 4D fMRI using hybrid 3DCNN and GRU networks** [12342-135]

SESSION 5 IMAGE SEGMENTATION

12342 1B **Subcortical brain segmentation with convolutional neural networks** [12342-32]

12342 1C **An improved U-shape neural network for soft exudate segmentation** [12342-101]

12342 1D **Hyperspectral remote sensing image semantic segmentation using extended extrema morphological profiles** [12342-13]

12342 1E **FSC-UNet: a lightweight medical image segmentation algorithm fused with skip connections** [12342-104]

12342 1F **Optic disc segmentation in retinal fundus images using improved CE-Net** [12342-30]

12342 1G **Super-resolution for semantic segmentation** [12342-16]

12342 1H **PDNet: an advanced architecture for polyp image segmentation** [12342-39]

12342 1I **Automatic heart segmentation based on convolutional networks using attention mechanism** [12342-38]

SESSION 6 IMAGE ANALYSIS AND METHODS

12342 1J **Generative image inpainting with residual texture prior and cross-layer contextual attention** [12342-167]

12342 1K **Metrical age assessment using image analysis and artificial neural networks** [12342-6]

12342 1L **Feature attention network (FA-Net): a deep-learning based approach for underwater single image enhancement** [12342-140]

12342 1M **Weakly supervised deep learning for cervical histopathology images analysis** [12342-90]

12342 1N **Surface defect sample generation method based on GAN** [12342-145]

12342 1O **Implementation of stereo matching algorithm based on Xavier edge computing platform** [12342-113]

12342 1P **Orthogonal matching pursuit with adaptive restriction for sparse subspace clustering** [12342-93]

12342 1Q **Mineralization information extraction of metal mining area based on Landsat images** [12342-46]

SESSION 7 IMAGE CLASSIFICATION AND ALGORITHMS

12342 1R **Comparison of different classification methods for autism spectrum diagnosis** [12342-127]

12342 1S **RHDDNet: multi-label classification-based detection of image hybrid distortions** [12342-45]

12342 1T **Improved video classification method based on non-parametric attention combined with self-supervision** [12342-18]

12342 1U **Compressing CNN by alternating constraint optimization framework** [12342-57]

12342 1V **Weakly-supervised attention mechanism via score-CAM for fine-grained visual classification** [12342-122]

12342 1W **Classification of fundus diseases based on meta-data and EB-IRV2 network** [12342-75]

12342 1X **Hyperspectral image classification based on dual-branch attention network with 3-D octave convolution** [12342-76]

12342 1Y **Fine-grained classification of natural scene grape images based on deep convolution networks** [12342-71]

12342 1Z **Multi-visual information fusion and aggregation for video action classification** [12342-96]

Part Two

SESSION 8 VIDEO ENCODING AND PROCESSING

12342 20 **Hardware-friendly fast rate-distortion optimized quantization algorithm for AVS3** [12342-5]

12342 21 **Multi-modal transformer for video retrieval using improved sentence embeddings** [12342-58]

12342 22 **Deep learning video analytics solutions for ocean surveillance systems** [12342-23]

12342 23 **Texture based adaptive computational resource allocation for fast AVS3 inter coding** [12342-88]

12342 24 **A foreground detection based video stabilization method and its application in aerospace measurement and control** [12342-146]

12342 25 **Video stabilization based on GMS and warping transform** [12342-91]

12342 26 **Semi-supervised multiple image transformation network for metastatic lymph node diagnosis from CT** [12342-157]

12342 27 **No-reference stereoscopic video quality assessment based on Tchebichef moment** [12342-160]

12342 28 **A hardware architecture of skip/direct mode for AVS3** [12342-8]

SESSION 9 MACHINE VISION AND APPLICATIONS

12342 29 **Real-time ranging of traffic signs for smart car environment perception** [12342-25]

12342 2A **End-to-end audiovisual feature fusion for active speaker detection** [12342-63]

12342 2B **An effective anti-interference visual tracking method** [12342-70]

12342 2C **AVR: attention based salient visual relationship detection** [12342-131]

12342 2D **Predicting good features using a hybrid feature for visual geolocation system** [12342-169]

12342 2E **Comparative analysis of feature-based methods and direct methods for unmanned system's vision navigation** [12342-12]

12342 2F **Infrared and visual image fusion via iterative quadtree decomposition and Bézier interpolation** [12342-82]

12342 2G **Measuring system for elongation at break of cable insulation sheath based on machine vision** [12342-24]

SESSION 10 OPTICAL AND IMAGING SYSTEMS

12342 2H **Modelling and simulation of lateral jet infrared radiation based on the light of sight method** [12342-61]

12342 2I **Infrared and visible images fusion method based on unsupervised learning** [12342-68]

12342 2J **Study on aero optical effect for space infrared detection system** [12342-118]

12342 2K **Spectrogram-based speech enhancement by spatial attention generative adversarial networks** [12342-115]

12342 2L **Segmentation based lidar odometry and mapping** [12342-81]

12342 2M **LSTM-SENet: category decoding model with an attention mechanism** [12342-147]

12342 2N **Temperature imaging network based on swin transformer for TDLAS tomography** [12342-40]

12342 2O **Accurate neuroanatomy segmentation using 3D spatial and anatomical attention neural networks** [12342-126]

12342 2P **A NSST-based infrared and visible image fusion method focusing on luminance effect** [12342-69]

12342 2Q **Generative adversarial networks based on MLP** [12342-141]

12342 2R **Fusion of infrared and visible sensor images based on anisotropic diffusion and fast guided filter** [12342-143]

SESSION 11 IMAGE QUALITY AND IMAGE SECURITY

12342 2S **No-reference video quality assessment using data dimensionality reduction and attention-based pooling** [12342-60]

12342 2T **A new zero-watermarking algorithm based on deep learning** [12342-10]

12342 2U **Blind image quality assessment based on transformer** [12342-44]

12342 2V **To sharpen or not to sharpen? On the effect of sharpening filters applied to magnified images** [12342-166]

12342 2W **Coverless image steganography based on neural style transfer** [12342-11]

12342 2X **Quality evaluation of dried carrot obtained in different drying conditions using deep convolutional neural networks** [12342-171]

12342 2Y **Improvement of attention modules for image captioning using pixel-wise semantic information** [12342-162]

12342 2Z **An efficient image encryption scheme based on variable row-columns scrambling and selective block diffusion** [12342-159]

12342 30 **Self-supervision based super-resolution approach for light field refocused image** [12342-22]

12342 31 **Reversible data hiding simultaneously using substitution of MSB and compression of LSB of encrypted image** [12342-56]

SESSION 12 IMAGE TRANSFORMATION AND COMPUTATION

12342 32 **Real-time image distortion correction based on FPGA** [12342-130]

12342 33 **Haze removal using a hybrid convolutional sparse representation model** [12342-37]

12342 34 **Cluster-based point cloud attribute compression using inter prediction and graph Fourier transform** [12342-67]

12342 35 **An improved method of image enhancement based on fuzzy theory** [12342-125]

12342 36 **An algorithm for facial mask area repair based upon deep learning** [12342-92]

12342 37 **DWT denoising for multi-variate time series forecasting** [12342-114]

SESSION 13 DIGITAL IMAGE PROCESSING AND APPLICATIONS

12342 38 **Track initiation algorithm for bearing-only target tracking in complex background** [12342-42]

12342 39 **Makeup transfer model based on BeautyGAN** [12342-111]

12342 3A **Solar panel fault diagnosis method based on improved YOLOv3** [12342-79]

12342 3B **Indoor target tracking with deep learning-based YOLOv3 model** [12342-165]

12342 3C **Rapid identification of mature Xanthoceras sorbifolium bunge** [12342-134]

12342 3D **Monocular inertial indoor location algorithm considering point and line features** [12342-164]

12342 3E **Automated osteoporosis prediction system using artificial intelligence to calculate cortical thickness index from hip X-rays** [12342-137]

12342 3F **Attention-guided feature fusion network for crowd counting** [12342-7]

12342 3G **Deformable voxel grids for shape comparisons** [12342-174]

SESSION 14 COMPUTER GRAPHICS AND DIGITAL PHOTOGRAPHY

12342 3H **Correction of images projected on non-white surfaces based on deep neural network** [12342-87]

12342 3I **Identifying recaptured images using deep hybrid correlation network** [12342-9]

12342 3J **ECA-based generative adversarial network for multi-focus colour image fusion** [12342-52]

12342 3K **High-speed videogrammetric technique for displacement measurement of floating offshore wind turbine model** [12342-28]

12342 3L **A parallel change detection method for spatiotemporally multi-temporal *SAR* image based on *Q-Learning* and wavelet** [12342-151]

12342 3M **Dynamic three-dimensional measurement using convolution neural network and binocular structured light system** [12342-123]

12342 3N **Coordinate transformation and three-dimensional backprojection algorithm for multiple receiver synthetic aperture sonar motion compensation** [12342-29]

SESSION 15 COMPUTER AND ELECTRONIC ENGINEERING

12342 3O **Radio frequency interference suppression based on two-dimensional frequency domain notch for P-band ultra-wideband SAR** [12342-31]

12342 3P **Estimating lower extremity joint angles during gait using reduced number of sensors count via deep learning** [12342-59]

12342 3Q **Tobacco plant disease dataset** [12342-89]

12342 3R **An one-dimensional signal based object detection network for apnea and hypopnea locating** [12342-55]

12342 3S **A collaborative spectrum sensing algorithm for cognitive radio based on related vector machine** [12342-150]

12342 3T **Concepts encoding via knowledge-guided self-attention networks** [12342-117]

12342 3U **Low frequency ultra-wideband BSAR electromagnetic scattering characteristic** [12342-107]