2023 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW 2023)

Waikoloa, Hawaii, USA 3 – 7 January 2023



IEEE Catalog Number: CFP23B39-POD ISBN:

979-8-3503-2057-2

Copyright © 2023 by the Institute of Electrical and Electronics Engineers, Inc. All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

*** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.

 IEEE Catalog Number:
 CFP23B39-POD

 ISBN (Print-On-Demand):
 979-8-3503-2057-2

 ISBN (Online):
 979-8-3503-2056-5

ISSN: 2572-4398

Additional Copies of This Publication Are Available From:

Curran Associates, Inc 57 Morehouse Lane Red Hook, NY 12571 USA

Phone: (845) 758-0400 Fax: (845) 758-2633

E-mail: curran@proceedings.com Web: www.proceedings.com



2023 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW) WACVW 2023

Table of Contents

Real-World Surveillance: Applications and Challenges

Synthehicle: Multi-Vehicle Multi-Camera Tracking in Virtual Cities
Exploiting Inter-Pixel Correlations in Unsupervised Domain Adaptation for Semantic Segmentation
Inseop Chung (Seoul National University), Jayeon Yoo (Seoul National University), and Nojun Kwak (Seoul National University)
Mixture Domain Adaptation To Improve Semantic Segmentation in Real-World Surveillance
Multimodal Data Augmentation for Visual-Infrared Person ReID With Corrupted Data
Transformer Based Multi-Grained Features for Unsupervised Person Re-Identification
EMOD: Efficient Moving Object Detection via Image Eccentricity Analysis and Sparse Neural Networks
Xiaomin Li (Texas State University; Ford Motor Company), Ramin Nabati (Ford Motor Company), Kunjan Singh (Ford Motor Company), Enrique Corona (Ford Motor Company), Vangelis Metsis (Texas State University), and Armin Parchami (Ford Motor Company)

Knowledge-Based Visual Context-Aware Framework for Applications in Robotic Services	ow-Resolution Thermal Sensor-Guided Image Synthesis	60
Christof W. Čorsel (LIACS, Leiden Úniversity), Michel van Lier (TNO Intelligent Imaging), Leo Kampmeijer (TNO Intelligent Imaging), and Erwin M. Bakker (LIACS, Leiden University) Multi-View Target Transformation for Pedestrian Detection	Doosoo Chang (Korea Telecom R&D Center) and Bohyung Han (Seoul	70
Wei-Yu Lee (TELIN-IPI, Ghent University-imec), Ljubomir Jovanov (TELIN-IPI, Ghent University-imec), and Wilfried Philips (TELIN-IPI, Ghent University-imec) A Transformer-Based Late-Fusion Mechanism for Fine-Grained Object Recognition in Videos	Christof W. Corsel (LIACS, Leiden University), Michel van Lier (TNO Intelligent Imaging), Leo Kampmeijer (TNO Intelligent Imaging), Nicolas Boehrer (TNO Intelligent Imaging), and Erwin M. Bakker (LIACS,	79
Jannik Koch (Fraunhofer IOSB), Stefan Wolf (Fraunhofer IOSB; Vision and Future Lab, Karlsruhe Institute of Technology), and Jürgen Beyerer (Fraunhofer IOSB; Vision and Future Lab, Karlsruhe Institute of Technology; Fraunhofer Center for Machine Learning) Bringing Generalization to Deep Multi-View Pedestrian Detection	Wei-Yu Lee (TELIN-IPI, Ghent University-imec), Ljubomir Jovanov (TELIN-IPI, Ghent University-imec), and Wilfried Philips (TELIN-IPI,	90
Jeet Vora (CVIT, IIIT Hyderabad), Swetanjal Dutta (CVIT, IIIT Hyderabad), Kanishk Jain (CVIT, IIIT Hyderabad), Shyamgopal Karthik (University of Tubingen), and Vineet Gandhi (CVIT, IIIT Hyderabad) Meet-in-the-Middle: Multi-Scale Upsampling and Matching for Cross-Resolution Face Recognition	Jannik Koch (Fraunhofer IOSB), Stefan Wolf (Fraunhofer IOSB; Vision and Future Lab, Karlsruhe Institute of Technology), and Jürgen Beyerer (Fraunhofer IOSB; Vision and Future Lab, Karlsruhe Institute of	100
Recognition	Jeet Vora (CVIT, IIIT Hyderabad), Swetanjal Dutta (CVIT, IIIT Hyderabad), Kanishk Jain (CVIT, IIIT Hyderabad), Shyamgopal Karthik	110
Neelu Madan (Visual Analysis and Perception Lab, Aalborg University), Mia Sandra Nicole Siemon (Visual Analysis and Perception Lab, Aalborg University; Milestone Systems), Magnus Kaufmann Gjerde (Visual Analysis and Perception Lab, Aalborg University), Bastian Starup Petersson (Visual Analysis and Perception Lab, Aalborg University), Arijus Grotuzas (Visual Analysis and Perception Lab, Aalborg University), Malthe Aaholm Esbensen (Visual Analysis and Perception Lab, Aalborg University), Ivan Adriyanov Nikolov (Visual Analysis and Perception Lab, Aalborg University), Mark Philip Philipsen (Visual Analysis and Perception Lab, Aalborg University), Kamal Nasrollahi (Visual Analysis and Perception Lab, Aalborg University; Milestone Systems), and Thomas B. Moeslund (Visual Analysis and Perception Lab, Aalborg University) An Automated and Scalable ML Solution for Mapping Invasive Species: The Case of the Australian Tree Fern in Hawaiian Forests	ecognition Klemen Grm (University of Ljubljana), Berk Kemal Özata (Istanbul Technical University; ASELSAN Inc.), Vitomir Štruc (University of	120
Australian Tree Fern in Hawaiian Forests	Neelu Madan (Visual Analysis and Perception Lab, Aalborg University), Mia Sandra Nicole Siemon (Visual Analysis and Perception Lab, Aalborg University; Milestone Systems), Magnus Kaufmann Gjerde (Visual Analysis and Perception Lab, Aalborg University), Bastian Starup Petersson (Visual Analysis and Perception Lab, Aalborg University), Arijus Grotuzas (Visual Analysis and Perception Lab, Aalborg University), Malthe Aaholm Esbensen (Visual Analysis and Perception Lab, Aalborg University), Ivan Adriyanov Nikolov (Visual Analysis and Perception Lab, Aalborg University), Mark Philip Philipsen (Visual Analysis and Perception Lab, Aalborg University), Kamal Nasrollahi (Visual Analysis and Perception Lab, Aalborg University; Milestone Systems), and Thomas B. Moeslund (Visual Analysis and Perception Lab,	130
	ustralian Tree Fern in Hawaiian Forests	140

VDiSC: An Open Source Framework for Distributed Smart City Vision and Biometric Surveillance Networks	148
Discriminative Sampling of Proposals in Self-Supervised Transformers for Weakly Supervised Object Localization	155
UPAR Challenge: Pedestrian Attribute Recognition and Attribute-Based Person Retrieval — Dataset, Design, and Results	166
Improving the Detection of Small Oriented Objects in Aerial Images	176
Combining Photogrammetric Computer Vision and Semantic Segmentation for Fine-Grained Understanding of Coral Reef Growth Under Climate Change	186
DepthCue: Restoration of Underwater Images Using Monocular Depth as a Clue	196
An Efficient Approach for Underwater Image Improvement: Deblurring, Dehazing, and Color	206

SeaDroneSim: Simulation of Aerial Images for Detection of Objects Above Water	16
Sea Ice Classification With Dual-Polarized SAR Imagery: A Hierarchical Pipeline	24
YOLOv7-Sea: Object Detection of Maritime UAV Images Based on Improved YOLOv7	33
A Novel Framework To Evaluate and Train Object Detection Models for Real-Time Victims Search and Rescue at Sea With Autonomous Unmanned Aerial Systems Using High-Fidelity Dynamic Marine Simulation Environment	39
Sonar Image Composition for Semantic Segmentation Using Machine Learning	18
A Survey on the Deployability of Semantic Segmentation Networks for Fluvial Navigation	55

Benjamin Kiefer (University of Tuebingen), Matej Kristan (University of Ljubljana), Janez Perš (University of Ljubljana), Lojze Žust (University of Ljubljana), Fabio Poiesi (Fondazione Bruno Kessler), Fabio Andrade (University of South-Eastern Norway), Alexandre Bernardino (University of Lisbon), Matthew Dawkins (Kitware), Jenni Raitoharju (University of Jyvaskyla), Yitong Quan (University of Tuebingen), Adem Atmaca (University of Tuebingen), Timon Höfer (University of Tuebingen), Qiming Zhang (The University of Sydney), Yufei Xu (The University of Sydney), Jing Zhang (The University of Sydney), Dacheng Tao (The University of Sydney; JD Explore Academy), Lars Sommer (Fraunhofer IOSB), Raphael Spraul (Fraunhofer IOSB), Hangyue Zhao (Beijing University of Posts and Telecommunications), Hongpu Zhang (Beijing University of Posts and Telecommunications), Yanyun Zhao (Beijing University of Posts and Telecommunications), Jan Lukas Augustin (Helmut Schmidt University), Eui-ik Jeon (University of Seoul), Impyeong Lee (University of Seoul), Luca Zedda (University of Cagliari), Andrea Loddo (University of Cagliari), Cecilia Di Ruberto (University of Cagliari), Sagar Verma (Granular AI), Siddharth Gupta (Granular AI), Shishir Muralidhara (TU Kaiserslautern), Niharika Hegde (TU Kaiserslautern), Daitao Xing (New York University), Nikolaos Evangeliou (New York University Abu Dhabi), Anthony Tzes (New York University Abu Dhabi), Vojtěch Bartl (Brno University of Technology), Jakub Špaňhel (Brno University of Technology), Adam Herout (Brno University of Technology), Neelanjan Bhowmik (Durham University), Toby P. Breckon (Durham University), Shivanand Kundargi (KLE Technological University), Tejas Anvekar (KLE Technological University), Ramesh Ashok Tabib (KLE Technological University), Uma Mudenagudi (KLE Technological University), Arpita Vats (Santa Clara University), Yang Song (Beijing University of Posts and Telecommunications), Delong Liu (Beijing University of Posts and Telecommunications), Yonglin Li (National University of Defense Technology), Shuman Li (National University of Defense Technology), Chenhao Tan (National University of Defense Technology), Long Lan (National University of Defense Technology), Vladimir Somers (EPFL; UCLouvain; Sportradar), Christophe De Vleeschouwer (EUCLouvain), Alexandre Alahi (EPFL), Hsiang-Wei Huang (University of Washington), Cheng-Yen Yang (University of Washington), Jeng-Neng Hwang (University of Washington), Pyong-Kun Kim (Electronics and Telecommunications Research Institute), Kwangju Kim (Electronics and Telecommunications Research Institute), Kyoungoh Lee (Electronics and Telecommunications Research Institute), Shuai Jiang (Beijing University of Posts and Telecommunications), Haiwen Li (Beijing University of Posts and Telecommunications), Zheng Ziqiang (Hong Kong University of Science and Technology), Tuan-Anh Vu (Hong Kong University of Science and Technology), Hai Nguyen-Truong (Hong Kong University of Science and Technology), Sai-Kit Yeung (Hong Kong University of Science and Technology), Zhuang Jia (Xiaomi Inc.), Sophia Yang (National Tsing Hua University), Chih-Chung Hsu (National Cheng Kung University), Xiu-Yu Hou (National Cheng Kung University), Yu-An Jhang (National Cheng Kung University), Simon Yang (National Taiwan Ocean University), and Mau-Tsuen Yang (National Dong-Hwa *University*)

Demographic Variations in Performance of Biometric Algorithms
The Gender Gap in Face Recognition Accuracy Is a Hairy Problem
The Impact of Racial Distribution in Training Data on Face Recognition Bias: A Closer Look 313 Manideep Kolla (HyperVerge Inc.) and Aravinth Savadamuthu (HyperVerge Inc.)
Multi-IVE: Privacy Enhancement of Multiple Soft-Biometrics in Face Embeddings
Analyzing the Impact of Gender Misclassification on Face Recognition Accuracy
Causal Structure Learning of Bias for Fair Affect Recognition
Pretraining Large Vision and Multimodal Models
Masked Autoencoder for Self-Supervised Pre-Training on Lidar Point Clouds

Benefits of Synthetically Pre-Trained Depth-Prediction Networks for Indoor/Outdoor Imag Classification	e 360
Kelly X. Lin (University of California, Santa Barbara), Irene Cho (University of California, Santa Barbara), Amey Walimbe (University of California, Santa Barbara), Bryan A. Zamora (University of California, Santa Barbara), Alex Rich (University of California, Santa Barbara), Sirius Z. Zhang (University of California, Santa Barbara), and Tobias Höllerer (University of California, Santa Barbara)	
RarePlanes Soar Higher: Self-Supervised Pretraining for Resource Constrained and Synthet Datasets Justin Downes (Amazon Web Services), Will Gleave (Amazon Web Services), and Dan Nakada (Amazon Web Services)	
MixGen: A New Multi-Modal Data Augmentation	379
Scalable and Accurate Self-Supervised Multimodal Representation Learning Without Align Video and Text Data	
(Amazon; UC Berkeley), Shalini Ghosh (Amazon), Anna Rumshisky (UMass Lowell; Amazon), and Wael Hamza (Amazon) Dealing With the Novelty in Open Worlds	
	401
Lowell; Amazon), and Wael Hamza (Amazon) Dealing With the Novelty in Open Worlds K-NN Embeded Space Conditioning for Enhanced Few-Shot Object Detection	
Lowell; Amazon), and Wael Hamza (Amazon) Dealing With the Novelty in Open Worlds K-NN Embeded Space Conditioning for Enhanced Few-Shot Object Detection	411
Dealing With the Novelty in Open Worlds K-NN Embeded Space Conditioning for Enhanced Few-Shot Object Detection	411

Phantom of Benchmark Dataset: Resolving Label Ambiguity Problem on Image Recognition in the Wild
Hyunhee Chung (SOCAR AI Research), Kyung Ho Park (SOCAR AI Research), Taewon Seo (SOCAR AI Research), and Sungwoo Cho (SOCAR AI Research)
Computer Vision for Winter Sports
Observation Centric and Central Distance Recovery for Athlete Tracking
Detecting Arbitrary Keypoints on Limbs and Skis With Sparse Partly Correct Segmentation
Masks
Fine-Grained Activity Recognition
Hand Guided High Resolution Feature Enhancement for Fine-Grained Atomic Action Segmentation Within Complex Human Assemblies
Video/Audio Quality in Computer Vision
On the Importance of Spatio-Temporal Learning for Video Quality Assessment
Accelerating AI Using Next-Generation Hardware: Possibilities and Challenges With Analog In-Memory Computing
Object-Ratio-Preserving Video Retargeting Framework Based on Segmentation and Inpainting 49 Jun-gyu Jin (Kyungpook National University), Jaehyun Bae (Kyungpook National University), Han-gyul Baek (Kyungpook National University), and Sang-hyo Park (Kyungpook National University)
Can Machines Learn To Map Creative Videos to Marketing Campaigns?

Face Image Quality Vector Assessment for Biometrics Applications 512 Nima Najafzadeh (West Virginia University), Hossein Kashiani (West Virginia University), Mohammad Saeed Ebrahimi Saadabadi (West Virginia University), Niloufar Alipour Talemi (West Virginia University), Sahar Rahimi Malakshan (West Virginia University), and Nasser M. Nasrabadi (West Virginia University)	
Face Image Quality for Actor Profile Image Curation	L
BAPose: Bottom-Up Pose Estimation With Disentangled Waterfall Representations	3
Image Quality Assessment Using Semi-Supervised Representation Learning	3
Non Homogeneous Realistic Single Image Dehazing	3
Subjective and Objective Video Quality Assessment of High Dynamic Range Sports Content 556 Zaixi Shang (Amazon Prime Video), Yixu Chen (Amazon Prime Video), Yongjun Wu (Amazon Prime Video), Hai Wei (Amazon Prime Video), and Sriram Sethuraman (Amazon Prime Video)	,
A Simple and Efficient Method for Dubbed Audio Sync Detection Using Compressive Sensing 568 Avijit Vajpayee (Amazon Prime Video), Zhikang Zhang (Arizona State University), Abhinav Jain (Amazon Prime Video), and Vimal Bhat (Amazon Prime Video)	;
Self-Supervised Effective Resolution Estimation With Adversarial Augmentations	3
Photorealistic Image and Environment Synthesis for Computer Vision	
CAD2Render: A Modular Toolkit for GPU-Accelerated Photorealistic Synthetic Data Generation for the Manufacturing Industry	3

Long-Range Recognition

Expanding Accurate Person Recognition to New Altitudes and Ranges: The BRIAR Dataset	593
Long Range Gait Matching Using 3D Body Fitting With Gait-Specific Motion Constraints	503
Attentive Sensing for Long-Range Face Recognition 6 Helio Perroni Filho (Centre for AI & Society, York University), Aleksander Trajcevski (Centre for AI & Society, York University), Kartikeya Bhargava (Centre for AI & Society, York University), Nizwa Javed (Centre for AI & Society, York University), and James H. Elder (Centre for AI & Society, York University)	513
Monocular Depth Estimation Challenge	
The Monocular Depth Estimation Challenge Jaime Spencer (University of Surrey), C. Stella Qian (Aston University), Chris Russell (Amazon), Simon Hadfield (University of Surrey), Erich Graf (University of Southampton), Wendy Adams (University of Southampton), Andrew J. Schofield (missing), James H. Elder (York University), Richard Bowden (University of Surrey), Heng Cong (Independent), Stefano Mattoccia (University of Bologna), Matteo Poggi (University of Bologna), Zeeshan Khan Suri (DENSO ADAS Engineering Services GmbH), Yang Tang (East China University of Science and Technology), Fabio Tosi (University of Bologna), Hao Wang (Independent), Youmin Zhang (University of Bologna), Yusheng Zhang (Independent), and Chaoqiang Zhao (East China University of Science and Technology)	623

Manipulation, Adversarial and Presentation Attacks in Biometrics
Face Forgery Detection Based on Facial Region Displacement Trajectory Series
Video Manipulations Beyond Faces: A Dataset With Human-Machine Analysis
COLFISPOOF: A New Database for Contactless Fingerprint Presentation Attack Detection Research
A Unified Model for Face Matching and Presentation Attack Detection Using an Ensemble of Vision Transformer Features
Explainable & Interpretable Artificial Intelligence for Biometrics
Learning Pairwise Interaction for Generalizable DeepFake Detection 672 Ying Xu (Norwegian University of Science and Technology), Kiran Raja (Norwegian University of Science and Technology), Luisa Verdoliva (University Federico II of Naples), and Marius Pedersen (Norwegian University of Science and Technology)
A Principal Component Analysis-Based Approach for Single Morphing Attack Detection
Finger-NestNet: Interpretable Fingerphoto Verification on Smartphone Using Deep Nested Residual Network
Human Saliency-Driven Patch-Based Matching for Interpretable Post-Mortem Iris Recognition 70: Aidan Boyd (University of Notre Dame), Daniel Moreira (University of Notre Dame), Andrey Kuehlkamp (University of Notre Dame), Kevin Bowyer

(University of Notre Dame), and Adam Czajka (University of Notre Dame)

Explainable Model-Agnostic Similarity and Confidence in Face Verification	711
Martin Knoche (Technical University of Munich), Torben Teepe	
(Technical University of Munich), Stefan Hörmann (Technical University	
of Munich), and Gerhard Rigoll (Technical University of Munich)	
	=4(
Author Index	719