

2022 18th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS 2022)

Madrid, Spain

29 November - 2 December 2022



IEEE Catalog Number: CFP22AVS-POD
ISBN: 978-1-6654-6383-6

**Copyright © 2022 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:	CFP22AVS-POD
ISBN (Print-On-Demand):	978-1-6654-6383-6
ISBN (Online):	978-1-6654-6382-9

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

CURRAN ASSOCIATES INC.
proceedings
.com

AVSS 2022

The 18th IEEE International Conference on Advanced Video and Signal-based Surveillance

Table of Contents

Surveillance Systems and Applications

01: Height Normalizing Image Transform for Efficient Scene Specific Pedestrian Detection	1
02: Accelerated Blind Deblurring Method via Video-based Estimation in Next Point Spread Functions for Surveillance	12
03: A Visual Surveillance System to Observe Realistic Road User Behavior for Improved Pedestrian and Cyclist Safety at Crossroads	20
04: UPM-GTI-Face: A dataset for the evaluation of the impact of distance and masks in face detection and recognition systems	28
05: Impact of Video Compression on the Performance of Object Detection Systems for Surveillance Applications	36
06: Robust Unseen Video Understanding for Various Surveillance Environments	44

Anomaly Detection and Recognition

07: Simultaneous Localization and Anomaly Detection from First-Person Video Data through a Coupled Dynamic Bayesian Network Model	52
08: Spatial-Temporal Transformer for Crime Recognition in Surveillance Videos	60
09: FPDM: Fisheye Panoptic segmentation dataset for Door Monitoring	68
10: Spatio-temporal predictive tasks for abnormal event detection in videos	76
11: Violence Detection using 3D Convolutional Neural Networks	84
12: Utilizing Crowd Collectiveness to Enhance Bottleneck Detection Based on the Lagrangian Framework	92

Biometrics and Person Re-Identification

13: FRIDA: Fisheye Re-Identification Dataset with Annotations	100
14: Fast High-Resolution Fingerprint Recognition using Domain-Knowledge Infused Global Descriptors	108
15: REET: Region Enhanced Transformer for Person Re- Identification	116
16: MASK-MORPH: Does Morphing of Custom 3D Face Masks Threatens the Face Recognition Systems?	124
17: Exploring attention on faces: similarities between humans and Transformers	132

18: Attribute De-biased Vision Transformer (AD-ViT) for Long-Term Person Re-identification	140
Detection and Tracking	
19: FB MOT: Flow Bridges the Gap between Detection and Tracking in Multiple Object Tracking	148
20: A Ship Detection Model Based on YOLOX with Lightweight Adaptive Channel Feature Fusion and Sparse Data Augmentation	155
21: Fusion Tree Network for RGBT Tracking	163
22: Tracking Failure Prediction for Siamese Trackers Based on Channel Feature Statistics	171
23: Refining Action Boundaries for One-stage Detection	179
24: Detection-aware multi-object tracking evaluation	187
Segmentation and Estimation	
25: Deep Learning and Computer Vision Techniques for Estimating Snow Coverage on Roads using Surveillance Cameras	195
26: Crowd counting Using DRL-based segmentation and RL- based density estimation	203
27: A Data-Driven Approach for the Localization of Interacting Agents via a Multi-Modal Dynamic Bayesian Network Framework	211
28: Dynamic Background Subtraction by Generative Neural Networks	219
29: Background Subtraction Network Module Ensemble for Background Scene Adaptation	227
30: REET: Graph Structure Learning Boosted Neural Network for Image Segmentation	235
Other Recognition Tasks in Computer Vision	
31: Deep Network for Extremely Low-Resolution Human Action Recognition	243
32: Reliability Scoring for the Recognition of Degraded License Plates	251
33: Learning Occlusion-Aware Dense Correspondences for Multi-Modal Images	259
34: Hand-Object Interaction Reasoning	267
35: Real time Vehicle Color Recognition on a budget: an investigation on the usage of CNN architectures	275
36: Customer Behavior Recognition Adaptable for Changing Targets in Retail Environments	283
Sensor Processing and Data Generation	
37: Robust Camera Calibration with a Single Image by Planar Homography	291
38: SAR Image Denoising in High Dynamic Range with Speckle and Thermal Noise Refinement Modeling	298
39: Dual Camera Based High Spatio-Temporal Resolution Video Generation For Wide Area Surveillance	306

40: Real Image Super-Resolution using GAN through modeling of LR and HR process	314
41: FB MOT: Consolidated Adversarial Network for Video De-raining and De-hazing	322
42: Inception-based Deep Learning Architecture for 3D Point Cloud Completion	330
Representation Learning	
43: Distance-dependent Feature Alignment and Selection for Imbalance 3D Point Cloud Object Detection	337
44: Deformable Modules for Flexible Feature Sampling on Vision Transformer	345
45: Temporal-Invariant Video Representation Learning with Dynamic Temporal Resolutions	353
46: An End-to-End Encoder-Decoder Network with Multi-scale Feature Pulling for Detecting Local Changes From Video Scene	361
47: RCRL: Replay-based Continual Representation Learning in Multi-task Super-Resolution	369
Audio-Visual Detection and Classification	
48: Detection of birds in a 3D environment referring to audio-visual information	384
49: Improved Separation of Closely-spaced Speakers by Exploiting Auxiliary Direction of Arrival Information within a U-Net Architecture	391
50: A Spatio-temporal Deep Learning Approach for Underwater Acoustic Signals Classification	399
51: Dual-valued Neural Networks	406
52: 3D Object Detection of Cars and Pedestrians by Deep Neural Networks from Unit-Sharing One-Shot NAS	414
53: An Experimental Study of the Concept Drift Challenge in Farm Intrusion Detection using Audio	422
DeepView Workshop	
54: Vision Transformers for Road Accident Detection from Dashboard Cameras	430
55: Statistical Analysis on Channel-wise Cross-correlation Features in Siamese Trackers	438
56: High-Dimensional Indexing Scheme for Scene Graph Retrieval	443
57: Automated Single Particle Growth Measurement using Segmentation	448