

# **2023 20th Conference on Robots and Vision (CRV 2023)**

**Montreal, Quebec, Canada  
6-8 June 2023**



**IEEE Catalog Number: CFP23347-POD**  
**ISBN: 979-8-3503-4140-9**

**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:	CFP23347-POD
ISBN (Print-On-Demand):	979-8-3503-4140-9
ISBN (Online):	979-8-3503-4139-3

**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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2023 20th Conference on Robots and Vision (CRV) **CRV 2023**

## Table of Contents

Welcome Message .....	x
Organizers .....	xi
Program Committee .....	xii
Sponsors .....	xiii

### Image Processing and Synthesis

LatentKeypointGAN: Controlling Images Via Latent Keypoints .....	1
<i>Xingzhe He (University of British Columbia, Canada), Bastian Wandt (University of British Columbia, Canada), and Helge Rhodin (University of British Columbia, Canada)</i>	
Towards Low-Cost Learning-based Camera ISP Via Unrolled Optimization .....	9
<i>Ali Mosleh (Huawei Noah's Ark Lab, Canada), Marzieh S. Tahaei (Huawei Noah's Ark Lab, Canada), James J. Clark (McGill University, Canada), and Vahid Partovi Nia (Huawei Noah's Ark Lab, Canada)</i>	

### SLAM

A Reversible Transformer for LiDAR Point Cloud Semantic Segmentation .....	19
<i>Perpetual Hope Akwensi (University of Calgary, Canada) and Ruisheng Wang (University of Calgary, Canada)</i>	
Continuous-Time Range-Only Pose Estimation .....	29
<i>Abhishek Goudar (University of Toronto, Canada; Vector Institute for Artificial Intelligence, Canada), Timothy D. Barfoot (University of Toronto, Canada; Vector Institute for Artificial Intelligence, Canada), and Angela P. Schoellig (University of Toronto, Canada; Vector Institute for Artificial Intelligence, Canada; Technical University of Munich, Germany)</i>	
Towards Open World NeRF-Based SLAM .....	37
<i>Daniil Lisus (University of Toronto, Canada), Connor Holmes (University of Toronto, Canada), and Steven Waslander (University of Toronto, Canada)</i>	

## Shapes

Sparse Shape Encoding for Topologically Improved Instance Segmentation .....	45
<i>Keyi Liu (York University, Toronto, Canada) and James H. Elder (York University, Toronto, Canada)</i>	
What Does the Occluding Contour Tell us About Quantitative Shape? .....	55
<i>Yiming Qian (Institute of High Performance Computing (IHPC) Agency for Science, Technology and Research (A*STAR), Singapore) and James H. Elder (York University, Canada)</i>	
Evaluating 3D Shape Analysis Methods for Robustness to Rotation Invariance .....	63
<i>Supriya Gadi Patil (Simon Fraser University), Angel X. Chang (Simon Fraser University), and Manolis Savva (Simon Fraser University)</i>	

## Object Detection and Tracking

InterTrack: Interaction Transformer for 3D Multi-Object Tracking .....	73
<i>John Willes (University of Toronto, Canada), Cody Reading (University of Toronto, Canada), and Steven L. Waslander (University of Toronto, Canada)</i>	
Along Similar Lines: Local Obstacle Avoidance for Long-Term Autonomous Path Following .....	81
<i>Jordy Sehn (University of Toronto, Canada), Yuchen Wu (University of Toronto, Canada), and Timothy D. Barfoot (University of Toronto, Canada)</i>	
aUToLights: A Robust Multi-Camera Traffic Light Detection and Tracking System .....	89
<i>Sean Wu (University of Toronto, Canada), Nicole Amenta (University of Toronto, Canada), Jiachen Zhou (University of Toronto, Canada), Sandro Papais (University of Toronto, Canada), and Jonathan Kelly (University of Toronto, Canada)</i>	

## Learning for Vision

Generalized Kronecker-based Adapters for Parameter-Efficient Fine-Tuning of Vision Transformers .....	97
<i>Ali Edalati (Huawei Noah's Ark Lab, Canada), Marawan Gamal Abdel Hameed (Université de Montréal, Canada), and Ali Mosleh (Huawei Noah's Ark Lab, Canada)</i>	
Learning-to-Count by Learning-to-Rank .....	105
<i>Adriano C. D'Alessandro (Simon Fraser University, Canada), Ali Mahdavi-Amiri (Simon Fraser University, Canada), and Ghassan Hamarneh (Simon Fraser University, Canada)</i>	
Kaskade-CNN and NB-KDE-CNN: CNN Architectures with KDE and Probabilistic Reasoning Layers .	113
<i>Michael Cormier (Mount Allison University, Canada)</i>	

## Segmentation and Classification

Class Instance Balanced Learning for Long-Tailed Classification .....	121
<i>Marc-Antoine Lavoie (University of Toronto, Canada) and Steven Waslander (University of Toronto, Canada)</i>	
Transformer-based Human Action Recognition with Dynamic Feature Selection .....	129
<i>Soufiane Lamghari (Polytechnique Montreal, Canada), Guillaume-Alexandre Bilodeau (Polytechnique Montreal, Canada), and Nicolas Saunier (Polytechnique Montreal, Canada)</i>	
ProPanDL: A Modular Architecture for Uncertainty-Aware Panoptic Segmentation .....	137
<i>Jacob Deery (University of Toronto, Canada), Chang Won Lee (University of Toronto, Canada), and Steven L. Waslander (University of Toronto, Canada)</i>	

## Poster Session

Contrastive Learning for Self-Supervised Pre-Training of Point Cloud Segmentation Networks With Image Data .....	145
<i>Andrej Janda (University of Toronto, Canada), Brandon Wagstaff (University of Toronto, Canada), Edwin G. Ng (University of Toronto, Canada), and Jonathan Kelly (University of Toronto, Canada)</i>	
Real-Time Instance Segmentation with Polygons using An Intersection-Over-Union Loss .....	153
<i>Katia Jodogne-Del Litto (Polytechnique Montréal, Canada) and Guillaume-Alexandre Bilodeau (Polytechnique Montréal, Canada)</i>	
Next-Best-View Selection for Robot Eye-in-Hand Calibration .....	161
<i>Jun Yang (University of Toronto, Canada), Jason Rebello (University of Toronto, Canada), and Steven L. Waslander (University of Toronto, Canada)</i>	
Diffusion Dataset Generation: Towards Closing the Sim2Real Gap for Pedestrian Detection .....	169
<i>Andrew Farley (Queen's University, Canada), Mohsen Zand (Queen's University, Canada), and Michael Greenspan (Queen's University, Canada)</i>	
Local Region-to-Region Mapping-based Approach to Classify Articulated Objects .....	177
<i>Ayush Aggarwal (University of Birmingham, United Kingdom), Rustam Stolkin (University of Birmingham, United Kingdom), and Naresh Marturi (University of Birmingham, United Kingdom)</i>	
Multi-Object Tracking and Segmentation with a Space-Time Memory Network .....	184
<i>Mehdi Miah (Polytechnique Montréal, Canada), Guillaume-Alexandre Bilodeau (Polytechnique Montréal, Canada), and Nicolas Saunier (Polytechnique Montréal, Canada)</i>	
Deformation Modeling for the Robotic Manipulation of 3D Elastic Objects using Physics-Informed Graph Neural Networks .....	194
<i>Angel J. Valencia (University of Ottawa, Canada) and Pierre Payeur (University of Ottawa, Canada)</i>	

Living in a Material World: Learning Material Properties from Full-Waveform Flash Lidar Data for Semantic Segmentation .....	202
<i>Andrej Janda (University of Toronto, Canada), Pierre Merriaux (Leddartech, Canada), Pierre Olivier (Leddartech, Canada), and Jonathan Kelly (University of Toronto, Canada)</i>	
Naive Scene Graphs : How Visual is Modern Visual Relationship Detection? .....	208
<i>David Abou Chacra (University of Waterloo, Canada) and John Zelek (University of Waterloo, Canada)</i>	
HyperMODEST: Self-Supervised 3D Object Detection with Confidence Score Filtering .....	217
<i>Jenny Xu (University of Toronto, Canada) and Steven L. Waslander (University of Toronto, Canada)</i>	
Adaptive Multiple Distributed Bidirectional Spiral Path Planning for Foraging Robot Swarms.....	225
<i>Qi Lu (The University of Texas Rio Grande Valley, USA) and Ryan Luna (The University of Texas Rio Grande Valley, USA)</i>	
Robust Scuba Diver Tracking and Recovery in Open Water Using YOLOv7, SORT, and Spiral Search .....	233
<i>Faraz Lotfi (McGill University, Canada), Khalil Virji (McGill University, Canada), and Gregory Dudek (McGill University, Canada)</i>	
CANDID: Correspondence AligNment for Deepburst Image Denoising .....	241
<i>Arijit Mallick (University of Tuebingen, Germany), Raphael Braun (University of Tuebingen, Germany), and Hendrik Pa Lensch (University of Tuebingen, Germany)</i>	
Hyperspectral Image Compression Using Implicit Neural Representations .....	248
<i>Shima Rezasoltani (University of Ontario Institute of Technology, Canada) and Faisal Z. Qureshi (University of Ontario Institute of Technology, Canada)</i>	
Enhancing Satellite Trail Detection in Night Sky Imagery with Automatic Salience Thresholding .....	256
<i>Nikolaus Kollo (Saint Mary's University, Canada), Yasushi Akiyama (Saint Mary's University, Canada), and Jiju Peethambaran (Saint Mary's University, Canada)</i>	
Empirical Thresholding on Spatio-Temporal Autoencoders Trained on Surveillance Videos in a Dementia Care Unit .....	265
<i>Shehroz S. Khan (KITE University Health Network, Canada; University of Toronto, Canada), Pratik K. Mishra (KITE University Health Network, Canada; University of Toronto, Canada), Bing Ye (KITE University Health Network, Canada), Kristine Newman (Toronto Metropolitan University, Canada), Andrea Iaboni (KITE University Health Network, Canada; University of Toronto, Canada), and Alex Mihailidis (KITE University Health Network, Canada; University of Toronto, Canada)</i>	
CrossMoCo: Multi-Modal Momentum Contrastive Learning for Point Cloud .....	273
<i>Sneha Paul (Concordia University, Canada), Zachary Patterson (Concordia University, Canada), and Nizar Bouguila (Concordia University, Canada)</i>	
Clarifying Myths About the Relationship Between Shape Bias, Accuracy, and Robustness .....	281
<i>Zahra Golpayegani (Concordia University, Canada), Patrick St-Amant (Zetane Systems Inc.), and Nizar Bouguila (Concordia University, Canada)</i>	

Rehabilitation Exercise Repetition Segmentation and Counting using Skeletal Body Joints .....	288
<i>Ali Abedi (University Health Network, Canada), Paritosh Bisht (Multimodal Perception Lab, International Institute of Information Technology Bangalore, India), Riddhi Chatterjee (Multimodal Perception Lab, International Institute of Information Technology Bangalore, India), Rachit Agrawal (Multimodal Perception Lab, International Institute of Information Technology Bangalore, India), Vyom Sharma (Multimodal Perception Lab, International Institute of Information Technology Bangalore, India), Dinesh Babu Jayagopi (Multimodal Perception Lab, International Institute of Information Technology Bangalore, India), and Shehroz S. Khan (University Health Network, Canada; University of Toronto, Canada)</i>	
Fast Fine-Tuning using Curriculum Domain Adaptation .....	296
<i>Lulan Shen (McGill University, Canada), Ibtihel Amara (McGill University, Canada), Ruofeng Li (McGill University, Canada), Brett Meyer (McGill University, Canada), Warren Gross (McGill University, Canada), and James J. Clark (McGill University, Canada)</i>	
Gradient-based Maximally Interfered Retrieval for Domain Incremental 3D Object Detection .....	304
<i>Barza Nisar (University of Toronto, Canada), Hruday Vishal Kanna Anand (University of Toronto, Canada), and Steven L. Waslander (University of Toronto, Canada)</i>	
Tree Health Assessment from UAV Images: Improving Object Detection and Classification using Hard Negative Mining and Semi-Supervised Autoencoder .....	312
<i>Hela Jemaa (University of Québec (TÉLUQ), Québec; Concordia University, Québec), Wassim Bouachir (University of Québec (TÉLUQ), Québec), Brigitte Leblon (Lakehead University, Ontario), Armand LaRocque (University of New Brunswick, New Brunswick), Ata Haddadi (A&amp;L Canada Laboratories, Ontario), and Nizar Bouguila (Concordia University, Québec)</i>	
Few-Shot Personality-Specific Image Captioning Via Meta-Learning .....	320
<i>Mehrdad Hosseinzadeh (University of Manitoba, Canada) and Yang Wang (Concordia University, Canada)</i>	
<b>Author Index .....</b>	<b>329</b>