Rawal **Khirodkar**

✉ rkhirodk@cs.cmu.edu | ⌂ rawalkhirodkar.github.io | ⌂ rawalkhirodkar

Research

My research is focused on developing computer vision algorithms to understand human interactions with other humans and the real world. My current projects span developing methods and infrastructure to obtain data helpful in estimating humans' 2D/3D poses under complex interaction scenarios. These projects are part of my longer-term goal of improving how we model human-to-human interactions in the real world. These interaction models can then simulate multi-human interactions and behavior in social settings.

Education

Ph.D., Robotics

Carnegie Mellon University, advised by Kris Kitani

GPA: 4.0/4.0

Masters, Robotics

Carnegie Mellon University, advised by Kris Kitani

- Graduated summa cum laude. GPA: 4.0/4.0
- Thesis: *Leveraging Simulation for Computer Vision*

B.Tech., Computer Science

Indian Institute of Technology, Bombay, advised by Ganesh Ramakrishnan

- Graduated magna cum laude. Overall GPA: 8.94/10
- Honors in Machine Learning, Minor in Mathematics
- Thesis: *Face Recognition in Videos*

Honors & Awards

2020	**Amazon PhD Fellowship**	*Pittsburgh, PA*
2019	**Government of India PhD Scholarship** - top 25 students in India	*Pittsburgh, PA*
2018	**Government of India MS Scholarship** - top 50 students in India	*Pittsburgh, PA*
2017	**IIT Bombay Student Teaching Award** - honorable mention	*Mumbai, India*
2013	**Indian National Physics Olympiad** - top 100 students in India	*Mumbai, India*
2009	**Indian National Talent Search Scholarship** - top 1% applicants	*Mumbai, India*

Publications

Observation-Centric SORT: Rethinking SORT for Robust Multi-Object Tracking

Jinkun Cao, Xinshuo Weng, Rawal Khirodkar, Jiangmiao Pang, Kris Kitani. *CVPR.* 2023

- OC-SORT is a simple, online and real-time multi-object tracker robust to occlusion and non-linear object motion.
- We address key drawbacks of the SORT framework by using an observation-centric perspective for tracking.
- We achieve state-of-the-art performance on datasets like MOT20, MOT17, KITTI and DanceTrack.

Sequential Ensembling for Semantic Segmentation

Rawal Khirodkar, Brandon Smith, Amit Agrawal, Antonio Criminisi. *preprint.* 2022

- We provide a rigorous ensembling benchmark for semantic segmentation.
- We propose a learnable and parameter-efficient ensembling technique SEQ-ENS which outperforms vanilla ensembling.
- SEQ-ENS achieves state-of-the-art results on Cityscapes, ADE20k and Pascal-VOC datasets.

Occluded Human Mesh Recovery

Rawal Khirodkar, Shashank Tripathi, Kris Kitani. *CVPR.* 2022

- OCHMR is a top-down method for human mesh recovery under severe occlusion.
- We condition the mesh regressor on the body-centermaps during training and inference.
- We achieve state-of-the-art performance on datasets like 3DPW-PC, OCHuman and CrowdPose.

Multi-Instance Pose Networks: Rethinking Top-Down Pose Estimation

Rawal Khirodkar, Visesh Chari, Amit Agrawal, Ambrish Tyagi. *ICCV.* 2021

- MIPNet is a fundamental change to top-down human pose estimation, predicting multiple pose instances given the input.
- The architecture is parameter efficient adding less than 1% parameters to the network.
- We achieve state-of-the-art performance on COCO and crowding datasets like OCHuman, CrowdPose.

RePOSE: Fast 6D Object Pose Refinement via Deep Texture Rendering

Shun Iwase, Xingyu Liu, Rawal Khirodkar, Rio Yokota, Kris Kitani. *ICCV.* 2021

- RePOSE uses object appearance along with geometric information for 6 D object pose estimation using a deep feature renderer.
- RePOSE is 3 times faster than existing approaches.
- We achieve state-of-the-art performance on LineMOD and Occlusion LineMOD datasets.

Adversarial Domain Randomization

Rawal Khirodkar, Kris Kitani. *preprint.* 2019

- We present a theoretical perspective on the effectiveness of domain randomization and its comparison with domain adaptation.
- ADR is an adversarial algorithm that improves the sample efficiency of domain randomization.
- ADR outperforms DR for image classification, object detection, and depth estimation on CLEVR, Syn2Real, and VIRAT datasets.

Domain Randomization for Scene Specific Object Detection & Pose Estimation

Rawal Khirodkar, Donghyun Yoo, Kris Kitani. *WACV.* 2019

- We design a simulator using Unreal Engine, capable of generating accurate annotations like instance segmentation and 6DoF pose.
- We bridged the reality gap by randomizing lighting, textures, distractors, and shapes of the objects in the scene.
- Our model trained only using synthetic data outperforms models trained using limited real data.

Professional Experience

Meta Reality Labs *Redmond, WA*
RESEARCH INTERN, ADVISED BY MINH VO *May - Aug 2022*

(Ongoing) The project focuses on 3D human understanding in the wild from the ego-centric perspective using Aria glasses.

Amazon *Sunnyvale, CA*
RESEARCH INTERN, ADVISED BY ANTONIO CRIMINISI *May - Aug 2021*

Developed a novel algorithm as an alternative to ensembling that sets a new state-of-the-art for semantic segmentation.

Amazon *Sunnyvale, CA*
RESEARCH INTERN, ADVISED BY AMBRISH TYAGI *May - Aug 2020*

Removed a fundamental limitation of pose estimation, currently the state-of-the-art for pose estimation under occlusion and crowding.

Trexquant *Stamford, CT (virtual)*
TECHNICAL INTERN, ADVISED BY TYGER PARK *May - Aug 2017*

Implemented an attention autoencoder for 34% data compression, the strategy was deployed into live trading in European markets.

Schlumberger *Mumbai, India*
RESEARCH INTERN, ADVISED BY SHUBHAM MISHRA *Jan - Mar 2016*

Augmented state-of-the-art oil field simulators with learning models, resulting in 40% speedup in latency.

Samsung *Bangalore, India*
RESEARCH INTERN, ADVISED BY VRAJESH SEJPAL *May - July 2016*

Developed a lexical parser and compiler for Bixby and was introduced in Samsung's flagship phone Galaxy S8.

Autodesk *Pune, India*
TECHNICAL INTERN, ADVISED BY MANISH AGRAWAL *May - July 2015*

Contributed to 123D Design (now Fusion 360) iOS app facilitating 2D deconstruction and reconstruction of 3D mesh models.

Teaching

16-720: Computer Vision	Fall 2021, CMU	**10-601: Machine Learning**	Fall 2018, CMU
16-720: Computer Vision	Fall 2020, CMU	**16-720: Computer Vision**	Spring 2018, CMU
16-811: Math Fundamental for Robotics	Fall 2019, CMU	**CS-101: Computer Science**	Spring 2017, IITB
16-831: Statistical Techniques in Robotics	Spring 2019, CMU	**CS-101: Computer Science**	Fall 2016, IITB

Service

Conference Reviewer

Conference on Computer Vision and Pattern Recognition (CVPR: 2022, 2021, 2020, 2019), European Conference on Computer Vision (ECCV: 2022, 2020), International Conference on Computer Vision (ICCV: 2021, 2019), Neural Information Processing Systems (NeurIPS: 2022), Association for Advancement of Artificial Intelligence (AAAI: 2020), Winter Conference on Applications of Computer Vision (WACV: 2022, 2021, 2020, 2019), Asian Conference on Computer Vision (ACCV: 2020, 2018).

Viax Research Mentor

Meet weekly to virtually mentor undergraduate students from around the world on research projects related to computer vision.

Aastha Foundation Tutor

Tutor high school students when visiting India belonging to underrepresented groups who are considering a career science and engineering.

Relevant Coursework

CMU Coursework

Computer Vision, Machine Learning, Image Synthesis, Visual Learning and Recognition, Deep Reinforcement Learning, Statistics in Robotics, Math Fundamentals in Robotics.

IITB Coursework

Advanced Machine Learning, Operating Systems, Automata Theory, Compilers, Parallel Computing, Algorithms Analysis, Artifical Intelligence, Real & Complex Analysis

Skills

Programming	Python, C++/C, Matlab, Solidity
Frameworks	Pytorch, Tensorflow, Unreal Engine, Blender, Unity
Languages	English, Hindi, Marathi