Zubair Irshad

RESEARCH SCIENTIST

Silicon Valley, CA

muhammadzubairirshad@gmail.com | zubairirshad.com | zubair-irshad | zubair-irshad

Deep Learning · Computer Vision · Neural Fields · Diffusion Models · 3D Foundation Models

Education

Georgia Institute of Technology
Atlanta, GA

PhD in Mechanical Engineering, Specialization: Artificial Intelligence and Deep Learning — *Aug. 2019 - Dec. 2023*

- Advisor: Dr. Zsolt Kira, PI Robotics Perception and Learning Lab
- Thesis: Learning 3D Robotics Perception using Inductive Priors. [Thesis Link]

Georgia Institute of Technology
Atlanta, GA

M.S. in Mechanical Engineering, Specialization: Robotics — *Aug. 2019 - Dec. 2023*

- Relevant Coursework: Robotics, Deep Learning, Machine Learning, Computer Vision, Intro to Robotics Research Linear Controls, Math. Methods in Applied Sciences, Reinforcement Learning, Visual Perception.

GIK University of Science & Technology
Topi, PK

B.S in Mechanical Engineering — *Aug. 2011 - May. 2015*

- Graduated with Magna Cum Laude. Awarded Dean honors roll for 8 semesters.

Work Experience

Toyota Research Institute, Robotics
Los Altos, CA

Research Scientist — *Jan. 2024 - Present*

- Working in the Computer Vision team on 3D perception system for Robotics
- Managing various university collaborations including University of California Berkeley, CMU, and others.

Toyota Research Institute, Machine Learning Research
Los Altos, CA

Deep Learning and Robotics Research Intern — *May. 2021 - Aug. 2022*

- Innovated a Single-Shot Reconstruction and Pose Estimation System (Paper accepted to ICRA'22, US Patent applied.)
- Project: Neural Fields for few-view view synthesis of outdoor scenes (Paper accepted to ICCV'23, US Patent applied)
- Project: Implicit generalized 3D object understanding (Paper accepted to ECCV'22, US Patent applied)

SRI International
Princeton, NJ

Deep Learning Research Intern — *May. 2020 - Aug. 2020*

- Project: Semantically-aware spatio-temporal agent for Vision-and-language navigation (Paper accepted to ICPR'22)

Georgia Institute of Technology
Atlanta, GA

Graduate Research Assistant — *Jan. 2019 - Dec. 2023*

- Sponsor: Toyota Research Institute. NeRFs for efficient and generalizable 3D scene representation and reconstruction.
- Sponsor: DAPRA Lifelong Learning Machines (*L2M*). Developed imitation learning agents for DeepMind StarCraft2.

Research & Projects

Neural Radiance Fields for Self-Supervised 3D Representation Learning
Georgia Tech

PhD [arXiv | Project Page | Github] — *Spring. 2023*

- Innovated a 3D pretraining strategy based on NeRF and masked auto-encoders for 3D scene understanding.
- Improved performance on downstream tasks including achieving SOTA 3D object detection with minimal fine-tuning.
- Accepted to ECCV 2024 and CVPR Neural Rendering Intelligence Workshop, 2024

NeRFs for few-shot scene synthesis of outdoor scenes

Georgia Tech

PhD [arXiv | Project Page | Github | Video] *Fall. 2022*

- Proposed an image-conditional triplane representation and introduced a novel dataset for training few-shot NeRFs.
- Improved SOTA on 3-view synthesis by absolute PSNR improvement of 1.5. Paper accepted to ICCV'23

Implicit representations for 3D Shape, Appearance & 6D Pose Optimization

Toyota Research Institute

Research Intern [arXiv | Project Page | Github | Video] *Spring. 2022*

- Proposed a novel differentiable database of implicit shape and texture priors for 3D object understanding in the real world.
- Improved SOTA on 6D pose and size estimation by 8% ↑ and PSNR by 50% ↑. Paper accepted to ECCV'22.

Object-centric Holistic 3D Scene Understanding

Toyota Research Institute

Research Intern [arXiv | Project Page | Github | Video] *Summer. 2021*

- Proposed a novel single-shot method to reconstruct 3D shape and recover poses of novel object instances in the real world.
- Improved performance on 6D pose and size estimation by 12% with real-time inference (40 FPS ↑). Accepted to ICRA'22.

Neural Perception & Planning for Embodied AI

Georgia Tech

PhD [Project Page | Github | arXiv] *Nov. 2017*

- Proposed hierarchical method for vision-and-language navigation; achieves state-of-the-art (14% SR ↑ and 14% SPL ↑)
- Introduced a novel dataset for long-horizon and cross-modal visual control of embodied agents. Accepted to ICRA'21.

Semantic Cross-Modal Reasoning for Embodied AI

SRI International

Research Intern [arXiv | Patent | Video] *Summer. 2020*

- Proposed a transformer-inspired semantically-aware method for Vision-and-language navigation task in Pytorch.
- Improved success performance in unseen simulation environments by 22% ↑

Autonomous Navigation of Mobile Robots

Georgia Tech

PhD [Project Page | Github] *Summer. 2020*

- Developed algorithms for successfully navigating the turtle-bot robot to given waypoints using camera, lidar, and ROS.
- Completed a maze navigation task using Classification and ROS and demonstrated the algorithm on turtlebot3 robot

Environment Perception and Control for Autonomous Driving

Georgia Tech

PhD [Project Page | Github] *Summer. 2020*

- Developed a visual odometry system: Estimating autonomous vehicle trajectory using feature matching given set of posed images
- Implementated Longitudinal and Lateral control to autonomously navigate a car through a set of given way points

Selected Publications

[1] **M.Z.Irshad**, S. Zakharov, V.Guizilini, A.Gaidon, Z.ira, R.Ambrus, NeRF-MAE: Masked AutoEncoders for Self-Supervised 3D Representation Learning for Neural Radiance Fields, European Conference on Computer Vision, ECCV 2024

[2] T. Ikeda, S. Zakharov, T. Ko, **M.Z.Irshad**, R. Lee, K. Liu, R. Ambrus, K. Nishiwaki, DiffusionNOCS: Managing Symmetry and Uncertainty in Sim2Real Multi-Modal Category-level Pose Estimation, IEEE International Conference on Intelligent Robot and Systems, IROS 2024

[3] M. Lunayach, S. Zakharov, D. Chen, R. Ambrus, Z. Kira, **M. Z. Irshad**, FSD: Fast Self-Supervised Single RGB-D to Categorical 3D Objects, International Conference on Robotics and Automation, ICRA 2024

[4] V.Jaganathan, H. Huang, **M.Z.Irshad**, V. Jampani, A.Raj, Z.Kira ICE-G: Image Conditional Editing of 3D Gaussian Splats, CVPR Workshop on AI for Content Creation, CVPRW 2024

[5] **M.Z.Irshad**, Learning 3D Robotics Perception using Inductive Priors, Georgia Institute of Technology, 2023

[6] **M.Z.Irshad**, S. Zakharov, K.Liu, V.Guizilini, T.Kollar, A.Gaidon, Z.Kira, R.Ambrus, NeO 360: Neural Fields for Sparse View Synthesis of Outdoor Scenes, International Conference on Computer Vision, ICCV 2023

[7] N.Heppert, **M.Z.Irshad**, S. Zakharov, K.Liu, R.Ambrus, J.Bohg, A.Valada, T.Kollar, CARTO: Category and Joint Agnostic Reconstruction of Articulated Objects, Computer Vision and Pattern Recognition Conference, CVPR 2023

[8] **M.Z.Irshad**, S. Zakharov, R.Ambrus, T.Kollar, Z.ira, A.Gaidon, ShAPO: Implicit Representations for Multi-Object Shape, Appearance, and Pose Optimization, European Conference on Computer Vision, ECCV 2022

[9] **M.Z.Irshad**, T.Kollar, M.Laskey, K.Stone, Z.Kira, CenterSnap: Single-Shot Multi-Object 3D Shape Reconstruction and Categorical 6D Pose and Size Estimation, IEEE International Conference on Robotics and Automation, ICRA 2022

[10] **M.Z.Irshad**, N.Mithun, Z.Seymour, H.P.Chiu, S.Samarasekera, R.Kumar, SASRA: Semantically-aware Spatio-Temporal Reasoning Agent for Vision-and-Language Navigation, International Conference on Pattern Recognition, ICPR 2022

[11] **M.Z.Irshad**, C.Y.Ma, Z.Kira, Hierarchical Cross-Modal Agent for Robotics Vision-and-Language Navigation, IEEE International Conference on Robotics and Automation, ICRA 2021

Talks

Apr 2024 Facebook AI Research Reading Group. Topic: Towards 3D Foundation Models. [Presentation Link]

Mar 2024 Stanford's Computer Vision Class. Topic: Neural Fields in Vision and Beyond. [Presentation Link]

Jan 2024 Robot Perception Class at Stanford. Topic: Neural Fields in Robotics and beyond. [Presentation Link]

Jun 2023 3D Deep Learning Reading Group. Topic: Neural Fields in Robotics (Part 1 and 2). [Presentation Link]

Apr 2023 Cohere for AI. Topic: Learning Object-centric Neural 3D Scene Representations. [Talk Link]

Apr 2023 Georgia Tech's Deep Learning Class. Topic: Learning Object-centric Centric Neural 3D Scene Representations

Patent Applications

[1] M. Lunayach, S. Zakharov, D. Chen, R. Ambrus, Z. Kira **M.Z.Irshad** : Fast Self-Supervised Single Image to Categorical 3D Objects Machine Learning Model Training. **US Patent App. 18/585,908**

[2] N.Heppert, **M.Z.Irshad**, S. Zakharov, K.Liu, R.Ambrus, J.Bohg, T.Kollar. : Category and Joint Agnostic Reconstruction of Articulated Objects. **US Patent App. 18/441,589**

[3] **M.Z.Irshad**, S.Zakharov, R.Ambrus, V.Guizilini, A.Gaidon, R.Ambrus. NeO 360: Neural Fields for Sparse View Synthesis of Outdoor Scenes. **US Patent App. 18/487,956**

[4] **M.Z.Irshad**, S.Zakharov, R.Ambrus, A.Gaidon. Implicit Representations for Multi-Object Shape, Appearance and Pose optimization. **US Patent App. 17/868,614**

[5] **M.Z.Irshad**, T.Kollar, M.Laskey, K.Stone. System and method for Single-shot multi-object 3D shape reconstruction and categorical 6D pose and size estimation. **US Patent App.63/243,984**

[6] H.Chiu, Z.Seymour, N.C.Mithun, **M.Z.Irshad**, S.Samarasekera, R.Kumar, K.Thopalli. System and method for efficient visual navigation. **US Patent App. 63/126,981**

Teaching

Deep Learning CS7643 (Co-taught with Facebook AI): *Atlanta, GA*

Graduate Teaching Assistant, Georgia Institute of Technology *Spring. 2021*

- Hosting office hours and grading assignments.

Robotics ME 7757 *Atlanta, GA*

Teaching Practicum, Georgia Institute of Technology *Spring. 2021*

- Co-teaching 3 classes, designing homework and exam.

Skills

Deep Learning Pytorch, Tensorflow, Huggingface, AWS

Computer Vision 3D Detection, 6D pose estimation, Neural Fields (NeRF), RGB-D Vision, 3D Gaussian Splatting

Machine Learning Deep Learning, Supervised learning, Foundation Models, Self-Supervised Learning

Programming Python, C++, OpenCV, ROS, Habitat, Habitat, Matterport3D, Gibson

Open-Source Software

[1] **Awesome Implicit NeRF Robotics** [Github], 1200+ stars, 70+ forks.

[2] **CenterSnap (Single-Shot Pose and Shape)** [Github], 280+ stars, 45+ forks.

[3] **NeO 360 (Generalizable NeRF)** [Github], 200+ stars, 9 forks.

[4] **ShAPO (Implicit Pose, Shape and Appearance of Objects)** [Github], , 170+ stars, 10 forks.

[5] **Robo-VLN (Robotics Vision-and-Language Navigation** [Github], 60+ stars, 8 forks.

[6] **Articulated Object NeRF** [Github], 45+ stars, 3 forks.

Professional Activities

Organizer RoboNerF: 1st Workshop On Neural Fields In Robotics [Webpage/Program/Accepted Papers]

Reviewer ICLR'24, ICRA'24, CVPR'24

Reviewer NeurIPS'23, Siggraph'23

Reviewer CVPR ' 23, Neural Fields Workshop CVPR'23

Reviewer ECCV ' 22, ICCV'23

Reviewer ICRA ' 22, RA-L ' 22

Reviewer IROS ' 21, ICRA ' 21

Honors & Awards

Academic

2023 **IEEE International Conference on Computer Vision (ICCV) Doctoral Consortium**, *U.S.A*

2017 **Fulbright International Scholar**, for M.S at Georgia Tech *U.S.A*

2018 **ASME RICE Cullimore Scholar**, for M.S at Georgia Tech *Atlanta, GA*

Domestic

2015 **Distinction/Dean honors roll**, (all semesters) for outstanding academic achievement *Topi, PK*

2018 **1st Place**, Technology Ventures class competition among 12 teams at Georgia Tech *Atlanta, GA*

Advising

M.S Shiva Gantha, Georgia Tech

M.S Vishnu Jaganathan, Georgia Tech

Fellows Ahnaf Munir / Anas Zafar, Fatima Fellowship, supported by Huggingface

Intern Nick Heppert, Toyota Research Institute

M.S. Mayank Lunayach, Georgia Tech