Zubair Irshad

Machine Learning Research Scientist

(+1) 470-309-7995 | muhammadzubairirshad@gmail.com | zubairirshad.com | zubair-irshad | zubair-irshad

Deep Learning · 3D Perception · Neural Fields (NeRFs) · 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
- Thesis: Learning 3D Robotics Perception using Inductive Priors. [Thesis Link]

Georgia Institute of Technology — *Atlanta, GA*

M.S. in Mechanical Engineering — *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, Planning & Controls (Autonomous Systems).

GIK University of Science & Technology — *Topi, PK*

B.S in Robotics/Mechanical Engineering, magna cum laude — *Aug. 2011 - May. 2015*

Work Experience

Toyota Research Institute, Robotics — *Los Altos, CA*

Research Scientist — *Jan. 2022 - Aug. 2022*

- Working in the Computer Vision team on 3D perception system for Robotics
- Managing and involved in 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 Mesh Reconstruction and Category-Level 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 neural representations for 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 reasoning agent for Vision-and-language navigation (U.S. Patent applied, Paper accepted to ICPR'22)

Georgia Institute of Technology — *Atlanta, GA*

Graduate Research Assistant — *Jan. 2019 - Present*

- 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

3D Foundation models for indoor 3D scene understanding — *Georgia Tech*

PhD — *Spring. 2023*

- Proposing a foundation model based on masked auto-encoders for 3D scene understanding. In submission
- Improved performance on various downstream tasks including achieving SOTA 3D object detection with minimal fine-tuning.

NeRFs for few-shot scene synthesis of outdoor scenes — *Georgia Tech*

PhD [arxiv | project page | github | video] — *Fall. 2022*

- Proposed an image-conditional triplanar representation for few-shot NeRF. Introduced large-scale dataset for outdoor scenes.
- Improved SOTA on 3-view view synthesis by absolute PSNR and SSIM improvement of 1.5 and 0.11. 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 databse of implicit shape and texture priors for 3D novel object understanding in the real world.
- Improved SOTA on 6D pose and size estimaiton by 8% ↑ and PSNR by 50% ↑ with latent optimization. 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 complete 3D shape and recover pose and size of novel object instances in real-world.
- Improved performance on 6D pose and size estimation by 12% with fast and 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 a hierarchical method for robotics vision-and-language navigation ; achieves state-of-the-art (14% SR ↑ and 14% SPL ↑)
- Introduced a novel data-set for long-horizon and cross-modal perception-based 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 while avoiding obstacles 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 for a aautonomous driving: Estimating vehicle trajectory using feature matching given set of posed images
- Implementation of Longitudinal and Lateral control to autonomously navigate a car through a set of given way points

Selected Publications

C1 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

C2 **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

C3 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

C4 **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

C5 **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

C6 **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

C7 **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

C8 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"**, Preprint, ICRA 2024

Patent Applications

P1 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**

P2 **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**

P3 **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**

P4 **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**

P5 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 frameworks	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
Simulators/Programming	Python, C++, OpenCV, ROS, Habitat, Habitat, Matterport3D, Gibson

Open-Source Software

Awesome Implicit NeRF Robotics	(github.com/zubair-irshad/Awesome-Implicit-NeRF-Robotics)
CenterSnap (Single-Shot Pose and Shape)	(github.com/zubair-irshad/CenterSnap)
ShAPO (Implicit Pose, Shape and Appearance of Objects)	(github.com/zubair-irshad/shapo)
Robo-VLN (Robotics Vision-and-Language Navigation	(github.com/GT-RIPL/robo-vln)
Articulated Object NeRF	(https://github.com/zubair-irshad/articulated-object-nerf)

Professional Activities

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

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