

# Zubair Irshad

PhD, Research Scientist

Silicon Valley, CA

Website: [zubairirshad.com](http://zubairirshad.com) | LinkedIn: [linkedin.com/in/zubair-irshad](https://www.linkedin.com/in/zubair-irshad) | Google Scholar: [Scholar](#)  
Github: [github.com/zubair-irshad](https://github.com/zubair-irshad) | Email: [muhammadzubairirshad@gmail.com](mailto:muhammadzubairirshad@gmail.com) | Contact: +1 470-309-7995

## Education

### PhD - Georgia Institute of Technology

Atlanta, GA

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. [\[Thesis Link, Talk Link\]](#)

### MS - Georgia Institute of Technology

Atlanta, GA

Mechanical Engineering, Specialization: Robotics

Aug. 2019 - Dec. 2023

- CGPA: 3.76/4.0. Relevant Coursework: Robotics, Deep Learning, Machine Learning, Computer Vision, Robotics Research, Interactive Robot Learning, Math. Methods in Applied Sciences, Reinforcement Learning, Visual Perception.

### BS - GIK University of Science & Technology

Topi, PK

Mechanical Engineering

Aug. 2011 - May. 2015

- CGPA: 3.76/4.0. Graduated with Magna Cum Laude. Awarded Dean honors roll for 8 semesters.

## Work Experience

### Research Scientist, Toyota Research Institute

Los Altos, CA

Developing cutting-edge vision and robotics technologies [\[7 publications, 3 patents\]](#)

Jan. 2024 - Present

- Leading various projects on 3D foundation models, Generative AI for Robotics and Multimodal AI.
- Managing various university collaborations including University of California Berkeley, CMU, and others.
- Publications accepted to CORL, ECCV and IROS. Applied 3 U.S. Patent applicants.

### Research Intern, Toyota Research Institute

Los Altos, CA

Led various machine learning vision projects across 3 internships [\[3 publications, 3 patents\]](#)

May. 2021 - Aug. 2022

- Innovated single-shot reconstruction and pose estimation (2 papers accepted to ICRA'22, ECCV'22. U.S. patents applied)
- Collaborated with team on fast articulated 3D object reconstruction (Paper accepted to CVPR'22, US Patent applied)

### Research Intern, SRI International

Princeton, NJ

Lead a project in the Computer Vision team [\[1 publication, 1 patent\]](#)

May. 2020 - Aug. 2020

- Innovated a novel architecture using semantics and spatiotemporal awareness for SOTA Vision-language navigation agent.

### Graduate Research Assistant, Georgia Institute of Technology

Atlanta, GA

Successfully led various robotics and machine learning projects [\[5 publications, 1 patent\]](#)

Jan. 2019 - Dec. 2023

- Sponsors: Toyota Research Institute, DAPRA Lifelong Learning Machines (L2M) and Georgia Tech.
- Led projects on Multimodal AI, Vision-Language, Gaussian Splat Editing, Shape Reconstruction, and 6D pose estimation. Research accepted at ICRA, ICCV, CVPRW.

## Research Projects

### Generative AI for Robotics

Toyota Research Institute

Research Scientist (2 publications) [\[RoVi-Aug\]](#) [\[DiffusionNOCS\]](#)

Spring. 2024 - Fall 2024.

- Collaborated with University of California Berkeley on Zero-shot viewpoint and cross-embodiment aware robot learning.
- Cross-collaboration with Woven by Toyota on improving symmetric object 6D pose estimation using Diffusion models.
- 2 papers accepted to CORL 2024 and IROS 2024. Press Coverage by [TechXplore](#).

## Foundation Model Distillation for 3D Navigation and Robotics Manipulation

Toyota Research Institute

Research Scientist (2 publications) [[POGS](#) | [LEGS](#)]

Fall. 2024 - Spring 2025.

- Collaborated with University of California Berkeley on two projects for foundation model distillation into 3D neural fields.
- Semantic foundation distillation enables text-driven navigation and zero-shot object tracking and manipulation.
- 2 papers accepted to IROS 2024 and ICRA 2024.

## Neural Radiance Fields for Self-Supervised and Generalizable 3D Representations

Georgia Tech / TRI

PhD & Research Scientist (3 publications) [[NeRF-MAE](#) | [NeO 360](#) | [MVG D](#)]

Spring. 2023. Spring 2025

- Zeros shot novel view and depth synthesis model trained on 60M multi-view image dataset.
- Innovated a 3D pretraining strategy based on NeRF and masked auto-encoders for 3D scene understanding.
- Improved SOTA on few-view synthesis for unbounded scenes. Proposed an image-conditional triplane representation.
- 2 papers accepted to ECCV 2024 and ICCV'23. U.S. patents applied.

## Object-centric 3D Pose, Shape and Appearance Reconstruction

Georgia Tech

PhD & Research Intern (4 publications) [[CenterSnap](#) | [ShAPO](#) | [FSD](#) | [CARTO](#)]

Summer 2021 - Spring. 2022

- 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 with real-time inference and self-supervised capabilities
- 3 papers accepted to ICRA 2024, ECCV 2022 and ICRA'22. U.S. patents applied.

## Multimodal AI for Embodied Semantic Perception & Planning

Georgia Tech/TRI

PhD & Research Intern (2 publications) [[Robo-VLN](#) | [SASRA](#)]

Fall 2020 & Spring 2021

- Proposed hierarchical & semantic transformer for vision-and-language navigation; achieves state-of-the-art (14% SR ↑)
- 2 papers accepted to ICRA'21 and ICPR'22. U.S. patent applied.

## Publications

---

- [1] **M.Z. Irshad**, S. Zakharov, V. Guizilini, A. Gaidon, Z. Kira, R. Ambrus, NeRF-MAE: Masked AutoEncoders for Self-Supervised 3D Representation Learning for Neural Radiance Fields, [European Conference on Computer Vision, ECCV 2024](#)
- [2] V. Guizilini, **M.Z. Irshad**, D. Chen, G. Shakhnarovich, R. Ambrus, Zero-Shot Novel View and Depth Synthesis with Multi-View Geometric Diffusion, [In Submission, 2025](#)
- [3] **M.Z. Irshad**, Mauro Comi, Yen-Chen Lin, Nick Heppert, Abhinav Valada, Zsolt Kira, Rares Ambrus, Jonathan Tremblay, Neural Fields in Robotics: A Survey, [In Submission, 2025](#)
- [4] J. Yu\*, K. Hari\*, K. El-Refai\*, A. Dalal, J. Kerr, C. M. Kim, R. Cheng, **M.Z. Irshad**, K. Goldberg, POGS: Persistent Object Gaussian Splat for Tracking Human and Robot Manipulation of Irregularly Shaped Objects, [International Conference on Robotics and Automation, ICRA 2025](#)
- [5] L. Chen\*, C. Xu\*, K. Dharmarajan, **M.Z. Irshad**, R. Cheng, K. Keutzer, M. Tomizuka, Q. Vuong, K. Goldberg, RoVi-Aug: Robot and Viewpoint Augmentation for Cross-Embodiment Robot Learning (**Oral Top 4.3%**), [Conference on Robot Learning, CORL 2024](#)
- [6] 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](#)
- [7] J. Yu\*, K. Hari\*, K. Srinivas\*, A. Rashid, C. M. Kim, J. Kerr, R. Cheng, **M.Z. Irshad**, A. Balakrishna, T. Kollar, K. Goldberg, LEGS: Incrementally Building Room-Scale Language-Embedded Gaussian Splats with a Mobile Robot, [IEEE International Conference on Intelligent Robot and Systems, IROS 2024](#)
- [8] J. Yu\*, T. Sadjadpour\*, A. O'Neill, M. Khfifi, L.Y. Chen, R. Cheng, **M.Z. Irshad**, A. Balakrishna, T. Kollar, K. Goldberg, MANIP: A Modular Architecture for Integrating Interactive Perception for Robot Manipulation, [IEEE International Conference on Intelligent Robot and Systems, IROS 2024](#)
- [9] 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](#)

- [10] **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](#)
- [11] 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](#)
- [12] **M.Z. Irshad\***, S. Zakharov\*, R. Ambrus, T. Kollar, Z. Kira, A. Gaidon, ShAPO: Implicit Representations for Multi-Object Shape, Appearance, and Pose Optimization, [European Conference on Computer Vision, ECCV 2022](#)
- [13] **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](#)
- [14] **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](#)
- [15] **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](#)

## Thesis, Workshop, and Symposium Publications

---

- [1] **M.Z. Irshad**, Learning 3D Robotics Perception using Inductive Priors, [Doctoral Dissertation, Georgia Institute of Technology, 2023](#)
- [2] **M.Z. Irshad**, S. Zakharov, V. Guizilini, A. Gaidon, Z. Kira, R. Ambrus, NeRF-MAE: Masked AutoEncoders for Self-Supervised 3D Representation Learning for Neural Radiance Fields, [CVPR Neural Rendering Intelligence Workshop, 2024 and ECCV Scalable 3D Scene Generation and Geometric Scene Understanding, 2024](#)
- [3] T. Ikeda, S. Zakharov, T. Ko, **M.Z. Irshad**, R. Lee, K. Liu, R. Ambrus, K. Nishiwaki, Handling Symmetry and Uncertainty in Category-level Pose Estimation with Diffusion Models, [ECCV Workshop on Recovering 6D Object Pose, 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**, S. Zakharov, R. Ambrus, T. Kollar, Z. Kira, A. Gaidon, ShAPO: Implicit Representations for Multi-Object Shape, Appearance, and Pose Optimization, [Baylearn Machine Learning Symposium 2022](#)

## Patent Applications

---

- [1] **M.Z. Irshad**, S. Zakharov, V. Guizilini, A. Gaidon, Z. Kira, R. Ambrus, Performing a three-dimensional Computer Vision task using a Neural Radiance Field grid representation of a scene produced from two-dimensional images of at least a portion of The scene. [US Patent App. 19/010,943](#)
- [2] V. Guizilini, **M.Z. Irshad**, D. Chen, R. Ambrus. Zero-Shot Novel View and Depth Synthesis with Multi-View Geometric Diffusion. [US Patent App. 63/737,994](#)
- [3] 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](#)
- [4] 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](#)
- [5] **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](#)
- [6] **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](#)
- [7] **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](#)
- [8] 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](#)

## Talks

---

**Dec 2024** Woven by Toyota (Tokyo, Japan) **Topic:** Learning 3D Robotics Perception using Inductive Priors.

**Aug 2024** Habib University (Karachi, PK) **Topic:** Towards Embodied 3D Foundation Models. [\[Presentation Link\]](#)

- Apr 2024** Facebook AI Research Reading Group (Bay Area, CA) **Topic:** Towards 3D Foundation Models. [[Presentation Link](#)]
- Mar 2024** Stanford's Computer Vision Class (Bay Area, CA) **Topic:** Neural Fields in Vision and Beyond. [[Presentation Link](#)]
- Jan 2024** Robot Perception Class at Stanford (Bay Area, CA) **Topic:** Neural Fields in Robotics and beyond. [[Presentation Link](#)]
- Jun 2023** 3D Deep Learning Reading Group (Online) **Topic:** Neural Fields in Robotics (Part 1 and 2). [[Presentation Link](#)]
- Apr 2023** Cohere for AI. **Topic:** Learning Object-centric Neural 3D Scene Representations (Online) [[Talk Link](#)]
- Apr 2023** Georgia Tech's Deep Learning Class (Atlanta, GA) **Topic:** Learning Object-centric Neural 3D Representations

## Professional Activities

---

- Organizer** Robo3D-VLMs: 3D Vision Language Models (VLMs) for Robotics Manipulation (CVPR'25): [[Webpage/Call for Papers](#)]
- Organizer** RoboNeRF: 1st Workshop On Neural Fields In Robotics (ICRA'24)[[Webpage/Program/Accepted Papers](#)]
- Reviewer** CVPR'25, ICLR'25
- Reviewer** ECCV'24, ICLR'24, ICRA'24, CVPR'24
- Reviewer** CVPR'23, Neural Fields Workshop CVPR'23, NeurIPS'23, Siggraph'23, ICCV'23
- Reviewer** ECCV ' 22, ICRA ' 22, RA-L ' 22
- Reviewer** IROS'21, ICRA'21

## Skills

---

- Deep Learning** Pytorch, Lightning AI, Tensorflow, Huggingface, AWS (Sagemaker)
- Computer Vision** 3D Detection, 6D pose estimation, Neural Fields ( NeRF), RGB-D Vision, 3D Gaussian Splatting
- Machine Learning** Supervised learning, Pretraining, Self-Supervised Learning, Few-shot Learning, Finetuning
- Programming** Python, C++, OpenCV, ROS, Linux, Github, LaTeX, Habitat, Habitat, Matterport3D, Gibson

## Teaching

---

### Deep Learning CS7643

Graduate Teaching Assistant, Georgia Institute of Technology (Co-taught with Facebook AI):

*Atlanta, GA*

*Spring. 2021*

- Hosting office hours and grading assignments.

### Robotics ME 7757

Teaching Practicum, Georgia Institute of Technology

*Atlanta, GA*

*Spring. 2021*

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

## Honors & Awards

---

- 2023 **IEEE International Conference on Computer Vision Doctoral Consortium**, ICCV *France*
- 2022 **Funding from Toyota Research Institute for my PhD research project**, PhD. at Georgia Tech *U.S.A*
- 2018 **ASME RICE Cullimore Scholar**, for M.S at Georgia Tech *Atlanta, GA*
- 2018 **1st Place**, Technology Ventures class competition among 12 teams at Georgia Tech *Atlanta, GA*
- 2017 **Fulbright International Scholar**, for M.S at Georgia Tech *U.S.A*
- 2015 **Distinction/Dean honors roll**, (all semesters) for outstanding academic achievement *Topi, PK*

## Open-Source Software

---

- [1] Awesome Implicit NeRF Robotics [[Github](#)], 1350+ stars, 80+ forks.
- [2] CenterSnap (Single-Shot Pose and Shape) [[Github](#)], 300+ stars, 45+ forks.
- [3] NeO 360 (Generalizable NeRF) [[Github](#)], 230+ stars, 10 forks.
- [4] ShAPO (Implicit Pose, Shape and Appearance of Objects) [[Github](#)], 180+ stars, 12 forks.

- [5] Robo-VLN (Robotics Vision-and-Language Navigation [\[Github\]](#), 60+ stars, 8 forks.
- [6] Awesome Robotics 3D [\[Github\]](#), 600+ stars, 30+ forks
- [6] NeRF-MAE (3D Representation Learning for NeRFs [\[Github\]](#), 85+ stars, 3 forks
- [7] Articulated Object NeRF [\[Github\]](#), 50 stars, 3 forks.

## Advising

---

- M.S** Gunjan Chablani, Georgia Tech – Now at Waymo
- PhD.** Xinan Zhang, Georgia Tech
- M.S** Shiva Gantha, Georgia Tech – Now R.E. at Matic
- M.S** Vishnu Jaganathan, Georgia Tech – Now at C3.AI
- Fellows** Ahnaf Munir / Anas Zafar, Fatima Fellowship, supported by Huggingface – Now PhD at UCF
- Intern** Nick Heppert, Toyota Research Institute — Now PhD at U.Frieburg
- M.S.** Mayank Lunayach, Georgia Tech – Now S.E at Google