

Zubair Irshad

PhD, Research Scientist

Silicon Valley, CA

Website: zubairirshad.com | LinkedIn: linkedin.com/in/zubair-irshad | Google Scholar: [Scholar](#)
Github: github.com/zubair-irshad | Email: 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

Working on Large Behavior Models, 3D Vision and Robotics [\[12 publications, 8 patents\]](#)

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 Models for Zero-shot 3D Navigation and Robotics Manipulation

Toyota Research Institute

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

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.
- Automatic post-hoc camera calibration of large-scale robotics dataset i.e. DROID. Released 36k calibrated episodes.
- 3 papers accepted to IROS 2024 and ICRA 2024 and RSS 2024.

Scaling-Up Camera Calibration for Large-scale Robot Learning Datasets

Toyota Research Institute

Research Scientist (2 publications) [[DROID](#) | [Open X-Embodiment](#)]

Fall. 2024 - Spring 2025.

- Publically released Posed-DROID; Provided around 36k good quality camera poses
- Utilized various foundation models for zero-shot post-hoc camera-calibration with robust quality metrics
- 2 papers accepted to RSS 2024, ICRA 2024. Best paper award at 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

- Zero-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 (5 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
- 4 papers accepted to CVPR 2025, 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, [Computer Vision and Pattern Recognition, CVPR 2025](#)
- [3] S. Iwase, **M.Z. Irshad**, K. Liu, V. Guizilini, R. Lee, T. Ikeda, A. Amma, K. Nishiwaki, K. Kitani, R. Ambrus, S. Zakharov, ZeroGrasp: Zero-Shot Shape Reconstruction Enabled Robotic Grasping, [Computer Vision and Pattern Recognition, CVPR 2025](#)
- [4] J. Yu, L. Fu, H. Huang, K. El-Refai, R. A. Ambrus, R. Cheng, **M.Z. Irshad**, K. Goldberg, Real2Render2Real: Scaling Robot Data Without Dynamics Simulation or Robot Hardware, [In Submission, 2025](#)
- [5] **M.Z. Irshad**, Mauro Comi, Yen-Chen Lin, Nick Heppert, Abhinav Valada, Zolt Kira, Rares Ambrus, Jonathan Tremblay, Neural Fields in Robotics: A Survey, [In Submission, 2025](#)
- [6] J. Li, H. Wang, **M.Z. Irshad**, I. Vasiljevic, M. R. Walter, V. Guizilini, G. Shakhnarovich, FastMap: Revisiting Dense and Scalable Structure from Motion, [In Submission, 2025](#)
- [7] 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](#)

- [8] 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](#)
- [9] Open X-Embodiment Collaboration. A. O'Neill, A. Rehman, A. Gupta,..., **M.Z. Irshad**, *et al.*, Open X-Embodiment: Robotic Learning Datasets and RT-X Models (**Best Paper Award.**), [IEEE International Conference on Robotics and Automation, ICRA 2024](#)
- [10] 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](#)
- [11] 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](#)
- [12] A. Khazatsky, K. Pertsch, S. Nair, A. Balakrishna,..., **M.Z. Irshad**, *et al.*, DROID: A Large-Scale In-The-Wild Robot Manipulation Dataset, [Robotics: Science and Systems, RSS 2024](#)
- [13] 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](#)
- [14] 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](#)
- [15] **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](#)
- [16] 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](#)
- [17] **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](#)
- [18] **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](#)
- [19] **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](#)
- [20] **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] X. Zhang, **M.Z. Irshad**, A. Yezzi, Y. Tsai, Z. Kira, EscherNet++: Simultaneous Amodal Completion and Scalable View Synthesis through Masked Fine-Tuning and Enhanced Feed-Forward 3D Reconstruction, [3rd CVPR Workshop on Generative Models for Computer Vision, 2025](#)
- [2] **M.Z. Irshad**, Learning 3D Robotics Perception using Inductive Priors, [Doctoral Dissertation, Georgia Institute of Technology, 2023](#)
- [3] **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](#)
- [4] 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](#)
- [5] 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](#)
- [6] **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] V. Guizilini, **M.Z.Irshad**, D. Chen, R. Ambrus. Systems and Methods for Generating a Scaled-Up and Fine-Tuned Diffusion Model for 3D Scene Reconstruction. [US Patent App. 19/187,140](#)
- [2] V. Guizilini, **M.Z.Irshad**, D. Chen, R. Ambrus. Systems and Methods for Scene Scale Normalization in Multi-View Depth Estimation. [US Patent App. 19/187,068](#)
- [3] S. Lin, J. Fang, **M.Z.Irshad**, V.C. Guizilini, R.A. Ambrus, G. Shakhnarovich, M. Walter. Systems and Methods for Reconstructing and Rendering Articulated Objects Using 3D Gaussian Splatting. [US Patent App. 19/187,006](#)
- [4] V. Guizilini, **M.Z.Irshad**, D. Chen, R. Ambrus. Multi-view Geometric Diffusion. [US Patent App. 19/184,534](#)
- [5] V. Guizilini, **M.Z.Irshad**, D. Chen, R. Ambrus. Multi-view Geometric using Incremental Conditioning. [US Patent App. 19/184,592](#)
- [6] **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](#)
- [7] 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](#)
- [8] 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](#)
- [9] 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](#)
- [10] **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](#)
- [11] **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](#)
- [12] **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](#)
- [13] H.Chui, 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	Learning 3D Robotics Perception using Inductive Priors , Woven by Toyota	<i>Tokyo, JP</i>
Aug 2024	Towards Embodied 3D Foundation Models , Habib University	<i>Karachi, PK</i>
Apr 2024	Towards 3D Foundation Models [Presentation Link], Facebook AI Research	<i>Bay Area, CA</i>
Mar 2024	Neural Fields in Vision and Beyond [Presentation Link], Stanford Computer Vision Class	<i>Bay Area, CA</i>
Jan 2024	Neural Fields in Robotics and beyond [Presentation Link], Sanford Robot Perception Class	<i>Bay Area, CA</i>
Jul 2023	Learning Object-Centric Neural 3D Scene Representations , Robotics and AI Institute	<i>Boston, MA</i>
Jun 2023	Neural Fields in Robotics [Presentation Link], 3D Deep Learning Reading Group	<i>Virtual</i>
Jun 2023	Learning Object-Centric Neural 3D Scene Representations , Qualcomm	<i>San Diego, CA</i>
Apr 2023	Learning Object-centric Neural 3D Scene Representations [Talk Link], Cohere for AI	<i>Virtual</i>
Apr 2023	Learning Object-centric Neural 3D Representations , Georgia Tech Deep Learning Class	<i>Atlanta, GA</i>

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, RSS'25, IROS'25, ICCV'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

Atlanta, GA

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

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.

Honors & Awards

2023 **IEEE International Conference on Computer Vision Doctoral Consortium**, ICCV

France

2022 **Funding from Toyota Research Institute for PhD**, 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] Open X-Embodiment [\[Github\]](#), 1.2k+ stars, 75+ forks.

[6] DROID Policy Learning [\[Github\]](#), 180+ stars, 16 forks.

[7] Robo-VLN (Robotics Vision-and-Language Navigation [\[Github\]](#), 60+ stars, 8 forks.

[8] Awesome Robotics 3D [\[Github\]](#), 600+ stars, 30+ forks

[9] NeRF-MAE (3D Representation Learning for NeRFs [\[Github\]](#), 100 stars, 3 forks

[10] 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