

Sahil Khose

[sahilkhose.github.io](https://github.com/sahilkhose) ◇ sahil.khose@gatech.edu ◇ [Google Scholar](#) ◇ (+1) 470 929 5628

RESEARCH INTERESTS

My research focuses on **multimodal learning**, particularly developing compact vision-language and audio-language models that integrate spatial, semantic, and temporal reasoning. I am also interested in **generalization under distribution shifts**, with prior work addressing domain and semantic gaps through synthetic-to-real transfer, zero-shot learning, and semi-supervised learning.

EDUCATION

Georgia Institute of Technology, Atlanta, USA

Ph.D. in Computer Science

2024 – Present

GPA: 4.0/4.0

Advisor: [Prof. Judy Hoffman](#)

Georgia Institute of Technology, Atlanta, USA

M.S. in Computer Science (ML specialization)

2022 – 2024

GPA: 4.0/4.0

Thesis: [Improving Real-World Aerial Scene Understanding with a Synthetic Dataset](#) [ECCV 2024]

Committee: [Prof. Judy Hoffman](#) (Advisor), [Prof. Zolt Kira](#), and [Prof. Humphrey Shi](#)

Manipal Institute of Technology, Manipal, India

B.Tech. in Computer and Communication Engineering

2018 – 2022

GPA: 8.56/10.0

Thesis: [Zero-Shot Domain Generalization: Unseen Classes in Unseen Domains](#)

PREPRINTS

P2. OVSG-VLM: Robust Open-Vocabulary 3D Scene Graph Generation with VLM

[Paper](#)

Under review at **NeurIPS 2025**

Mengqi Zhang, **Sahil Khose**, Fiona Ryan, Judy Hoffman

Developed a scalable, 7B open-source VLM that unifies spatial and semantic 3D scene graph generation, surpassing GPT-4o-based methods on both closed- and open-vocab 2D/3D SGG benchmarks.

P1. Beyond Single Modalities: Lightweight Joint-Training for Vision + Audio Generalist LLMs

[Paper](#)

Sahil Khose, Manushree Vasu, Humphrey Shi, Judy Hoffman

Designed a lightweight, jointly-trained 7B multimodal LLM that outperforms larger specialist and generalist models on vision and audio benchmarks by reducing cross-modal interference through simple MLP projectors.

CONFERENCE PAPERS

C4. SkyScenes: A Synthetic Dataset for Aerial Scene Understanding

European Conference on Computer Vision (**ECCV**) 2024

[Paper](#) | [Dataset](#) | [GitHub](#)

Sahil Khose*, Anisha Pal*, Aayushi Agarwal*, Deepanshi*, Judy Hoffman, Prithvijit Chattopadhyay

Built a replayable CARLA pipeline that systematically varies viewpoint, weather, and lighting to surface domain-shift failure modes, then leveraged it to boost real-world aerial segmentation via syn-to-real transfer.

C3. LatentDR: Improving Model Generalization With Sample-Aware Latent Degradation & Restoration

Winter Conference on Applications of Computer Vision (**WACV**) 2024

[Paper](#)

Ran Liu, **Sahil Khose**, Jingyun Xiao, Lakshmi Sathidevi, Keerthan Ramnath, Zolt Kira, Eva L. Dyer

A plug-and-play, sample-aware latent augmentation that lifts domain-generalization accuracy by up to 3 points on DomainBed and outperforms SoTA on medical and long-tail tasks.

C2. INDICON 2023: Explainable Classification of Macular Degeneration Using Deep Learning

[IEEE](#) | [Paper](#)

Sahil Khose*, Ankita Ghosh*, Yogish Kamath, Neetha Kuzhupilly, Harish Kumar J. R.

C1. INDICON 2023: Fovea Segmentation Using Semi-Supervised Learning

[IEEE](#) | [Paper](#)

Ankita Ghosh*, **Sahil Khose***, Yogish Kamath, Neetha Kuzhupilly, Harish Kumar J. R.

WORKSHOP PAPERS († for first author)

W7. NeurIPS 2022: Continual VQA for Disaster Response Systems † [Poster] Tackling Climate Change with ML at NeurIPS 2022	Sep 2022 GitHub Paper
W6. ICML 2022: An Efficient Modern Baseline for FloodNet VQA † [Best Paper Award] New in ML at ICML 2022	May 2022 GitHub Paper
W5. ACL 2022: Transformer based ensemble for emotion detection [Oral] WASSA at ACL 2022	Mar 2022 GitHub Paper
W4. NeurIPS 2021: A Studious Approach to Semi-Supervised Learning † [Poster] ICBINB at NeurIPS 2021	Sep 2021 GitHub Paper
W3. NeurIPS 2021: XCI-Sketch [Oral] New in ML, [Paper] ML4CD, [Paper] CtrlGen, [Poster] DGM at NeurIPS 2021	Aug 2021 GitHub Paper
W2. NeurIPS 2021: Semi-Supervised Classification & Segmentation on High Resolution Aerial Images † [Spotlight Paper] Tackling Climate Change with ML at NeurIPS 2021	May 2021 GitHub Paper
W1. NAACL 2021: BERT Transformers in Extraction of Health Information from Social Media † [Top Performer Award] Published in proceedings of NAACL 2021 at SMM4H workshop	Apr 2021 GitHub Paper

RESEARCH EXPERIENCE

Georgia Institute of Technology, Atlanta, USA Graduate Research Assistant at Hoffman AI Lab	Jan 2023 – Present Advisor – Prof. Judy Hoffman
<ul style="list-style-type: none">• OVSG-VLM: 3D Scene Graph Generation model for spatial and semantic reasoning in real-world robotics tasks. [P2]• Multimodality: Compact vision-audio LLM with MLP projectors and joint training, reduces cross-modal interference. [P1]• Syn-to-real: Developed syn-to-real adaptation to raise off-road semantic segmentation performance.• SkyScenes: Synthetic aerial benchmark that lifts model performance when transferring from sim-to-real. [C4] ECCV '24	
Georgia Institute of Technology, Atlanta, USA Graduate Research Assistant at Neural Data Science Lab (NerDS)	Spring 2023 Advisor – Prof. Eva Dyer
<ul style="list-style-type: none">• LatentDR: a plug-and-play module to counter diversity shift without architecture changes. [C3] WACV '24	
Indian Institute of Science, Bangalore, India AI Research Assistant at Artificial Intelligence and Robotics Lab	Jul 2021 – Jul 2022 Advisors – Prof. S. Sundaram & Dr. Chandan Gautam
<ul style="list-style-type: none">• Bachelor's Thesis: Jointly addressing domain shift + semantic shift to recognize unseen classes in unseen domains.	
Manipal Institute of Technology, Manipal, India Medical AI Research Assistant	Apr 2021 – Oct 2022 Advisor – Prof. Harish Kumar JR
<ul style="list-style-type: none">• Developed a medical diagnosis system for fovea segmentation using semi-supervised segmentation. [C2]• Designed a macular degeneration classification system with interpretability for ophthalmology diagnosis. [C1]	
Project MANAS – AI Robotics Research Team, MIT, Manipal, India AI Perception Developer GitLab Website	Sep 2018 – May 2021
<ul style="list-style-type: none">• World Rank 1 at IGVC 2019 (UGV) & winner of the Mahindra \$1M Challenge (out of 153 self-driving car teams).• Implemented Lane Detection, Speed Bump Detection, Driving Imitation System, Depth Map Generation.	

SELECTED PROJECTS

PR2. Domain Generalization: Tackling Diversity & Correlation Shifts YouTube GitHub	Fall 2022
<ul style="list-style-type: none">• Unified RSC and VREx to jointly mitigate diversity shift + spurious-correlation shift by equalizing cross-domain risk and suppressing shortcut cues (e.g., dominant colors/edges).• Established new SOTA on all six DomainBed datasets, with pronounced gains on color-biased gender-classification tasks.	
PR1. Zero-Shot Domain Generalization: Unseen Classes in Unseen Domains BTech Thesis	Spring 2022
<ul style="list-style-type: none">• Developed a CLIP-powered Class-Normalization Zero-Shot Learning framework that jointly addresses domain shift + semantic shift on DomainNet, enabling one model to recognize unseen classes in unseen domains.• Beat CuMix and DIN on all five held-out domains in ~30 s/train run and proposed a realistic DGZSL evaluation protocol.	

PROFESSIONAL SERVICE

Conference Reviewer: **NeurIPS 2025**, **CVPR 2025**, **ECCV 2024**
Workshop Reviewer: **CVPR-W 2025** (EMACS), **NeurIPS-W 2023** (ICBINB, DGM4H), **ICCV-W 2023** (WiCV), **NAACL-W 2021** (SMM4H)
Volunteer: **ICRA 2025** – Atlanta, GA, **NeurIPS 2022** – New Orleans, LA
Teaching Experience: **Graduate Teaching Assistant** for CS 7647 instructed by **Prof. Judy Hoffman** [Fall 2023]