Sahil Khose

sahilkhose.github.io ◇ 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 2024 – Present
Ph.D. in Computer Science *GPA: 4.0/4.0*
Advisor: *Prof. Judy Hoffman*

Georgia Institute of Technology, Atlanta, USA 2022 – 2024
M.S. in Computer Science (ML specialization) *GPA: 4.0/4.0*
Thesis*: Improving Real-World Aerial Scene Understanding with a Synthetic Dataset* [ECCV 2024]
Committee: *Prof. Judy Hoffman (Advisor), Prof. Zsolt Kira, and Prof. Humphrey Shi*

Manipal Institute of Technology, Manipal, India 2018 – 2022
B.Tech. in Computer and Communication Engineering *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
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
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, Zsolt 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 Kuzhuppilly, Harish Kumar J. R.*

C1. INDICON 2023: Fovea Segmentation Using Semi-Supervised Learning IEEE | Paper
Ankita Ghosh,* ***Sahil Khose****, Yogish Kamath, Neetha Kuzhuppilly, Harish Kumar J. R.*

Workshop Papers († for first author)

W7. NeurIPS 2022: Continual VQA for Disaster Response Systems † Sep 2022
[Poster] *Tackling Climate Change with ML at* ***NeurIPS 2022*** GitHub | Paper

W6. ICML 2022: An Efficient Modern Baseline for FloodNet VQA † May 2022
[Best Paper Award] *New in ML at* ***ICML 2022*** GitHub | Paper

W5. ACL 2022: Transformer based ensemble for emotion detection Mar 2022
[Oral] *WASSA at* ***ACL 2022*** GitHub | Paper

W4. NeurIPS 2021: A Studious Approach to Semi-Supervised Learning † Sep 2021
[Poster] *ICBINB at* ***NeurIPS 2021*** GitHub | Paper

W3. NeurIPS 2021: XCI-Sketch Aug 2021
[Oral] *New in ML,* ***[Paper]*** *ML4CD,* ***[Paper]*** *CtrlGen,* ***[Poster]*** *DGM at* ***NeurIPS 2021*** GitHub | Paper

W2. NeurIPS 2021: Semi-Supervised Classification & Segmentation on High Resolution Aerial Images † May 2021
[Spotlight Paper] *Tackling Climate Change with ML at* ***NeurIPS 2021*** GitHub | Paper

W1. NAACL 2021: BERT Transformers in Extraction of Health Information from Social Media † Apr 2021
[Top Performer Award] *Published in proceedings of* ***NAACL 2021*** *at SMM4H workshop* GitHub | Paper

Research Experience

Georgia Institute of Technology, Atlanta, USA Jan 2023 – Present
Graduate Research Assistant at ***Hoffman AI Lab*** *Advisor – Prof. Judy Hoffman*
- **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 Spring 2023
Graduate Research Assistant at ***Neural Data Science Lab*** *(NerDS)* *Advisor – Prof. Eva Dyer*
- **LatentDR**: a plug-and-play module to counter **diversity shift** without architecture changes. [C3] WACV '24

Indian Institute of Science, Bangalore, India Jul 2021 – Jul 2022
AI Research Assistant at ***Artificial Intelligence and Robotics Lab*** *Advisors – Prof. S. Sundaram & Dr. Chandan Gautam*
- Bachelor's Thesis: Jointly addressing **domain shift** + **semantic shift** to recognize unseen classes in unseen domains.

Manipal Institute of Technology, Manipal, India Apr 2021 – Oct 2022
Medical AI Research Assistant *Advisor – Prof. Harish Kumar JR*
- 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 Sep 2018 – May 2021
AI Perception Developer GitLab | Website
- **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
- 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
- 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]