

Sahil Khose

[sahilkhose.github.io](https://github.com/sahilkhose) ◇ sahilkhose18@gmail.com ◇ [Google Scholar](https://scholar.google.com/citations?user=sahilkhose) ◇ [in/sahilkhose](https://www.linkedin.com/in/sahilkhose) ◇ (+1) 470 929 5628

RESEARCH INTERESTS

My research advances **vision-language models (VLMs)** by extending their capabilities across modalities, spatial reasoning, and evaluation – integrating **audio**, enhancing **spatial understanding**, and enabling automatic evaluation of **world models** for robotic manipulation. I have also worked on **syn-to-real transfer** and **domain generalization**.

EDUCATION

Georgia Institute of Technology, Atlanta, USA <i>Ph.D. in Computer Science</i> <i>Advisor: Prof. Judy Hoffman</i>	2024 – Present GPA: 4.0/4.0
Georgia Institute of Technology, Atlanta, USA <i>M.S. in Computer Science (ML specialization)</i> <i>Thesis: Improving Real-World Aerial Scene Understanding with a Synthetic Dataset [ECCV 2024]</i> <i>Committee: Prof. Judy Hoffman (Advisor), Prof. Zolt Kira, and Prof. Humphrey Shi</i>	2022 – 2024 GPA: 4.0/4.0
Manipal Institute of Technology, Manipal, India <i>B.Tech. in Computer and Communication Engineering</i> <i>Thesis: Zero-Shot Domain Generalization: Unseen Classes in Unseen Domains</i>	2018 – 2022 GPA: 8.56/10.0

RESEARCH EXPERIENCE

Georgia Institute of Technology, Atlanta, USA <i>Graduate Research Assistant at Hoffman AI Lab</i> <ul style="list-style-type: none">WFM for robotic manipulation: Developing VLM-based benchmarks for automatic evaluation of world models. [P3]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 syn-to-real. [C4] ECCV '24	Jan 2023 – Present Advisor – Prof. Judy Hoffman
Georgia Institute of Technology, Atlanta, USA <i>Graduate Research Assistant at Neural Data Science Lab (NerDS)</i> <ul style="list-style-type: none">LatentDR: a plug-and-play module to counter diversity shift without architecture changes. [C3] WACV '24	Spring 2023 Advisor – Prof. Eva Dyer
Indian Institute of Science, Bangalore, India <i>AI Research Assistant at Artificial Intelligence and Robotics Lab</i> <ul style="list-style-type: none">Bachelor's Thesis: Jointly addressing domain shift + semantic shift to recognize unseen classes in unseen domains.	Jul 2021 – Jul 2022 Advisors – Prof. S. Sundaram & Dr. Chandan Gautam
Manipal Institute of Technology, Manipal, India <i>Medical AI Research Assistant</i> <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]	Apr 2021 – Oct 2022 Advisor – Prof. Harish Kumar JR
Project MANAS – AI Robotics Research Team, MIT, Manipal, India <i>AI Perception Developer</i> GitLab Website <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.	Sep 2018 – May 2021

ACHIEVEMENTS

• Best Paper Award at the New In ML workshop.	ICML 2022
• Spotlight Paper at the Tackling Climate Change with ML workshop.	NeurIPS 2021
• Top Performer Award and special recognition on multi-task performance at the SMM4H workshop.	NAACL 2021
• Project MANAS stood World Rank 1 at the 27th Intelligent Ground Vehicle Competition .	IGVC 2019
• IGVC 2019 Awards: Grand Award - 1st (Lescoe Cup), Interoperability - 1st, Design - 2nd, Cybersecurity - 3rd.	IGVC 2019
• Project MANAS won the the Mahindra \$1Million Challenge (top 13 out of 153 teams in India) .	2019

ONGOING

P3. VLMs for Robust Evaluation of World Models in Robotic Manipulation

In submission

Sahil Khose, Prithvijit Chattopadhyay, Judy Hoffman

Developing robust benchmarks for world models in robotic manipulation, designing VLM-based evaluators of task completion and object consistency to improve reliability and alignment with human judgments.

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

Under review

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

[Preprint](#)

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 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

W7. NeurIPS 2022: Continual VQA for Disaster Response Systems

Sep 2022

[Poster] Tackling Climate Change with ML at **NeurIPS 2022**

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

Aditya Kane*, V Manushree*, **Sahil Khose***

W6. ICML 2022: An Efficient Modern Baseline for FloodNet VQA

May 2022

[Best Paper Award] New in ML at **ICML 2022**

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

Aditya Kane*, **Sahil Khose***

W5. ACL 2022: Transformer based ensemble for emotion detection

Mar 2022

[Oral] WASSA at **ACL 2022**

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

Aditya Kane, Shantanu Patankar, **Sahil Khose**, Neeraja Kirtane

W4. NeurIPS 2021: A Studious Approach to Semi-Supervised Learning

Sep 2021

[Poster] ICBINB at **NeurIPS 2021**

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

Sahil Khose*, Shruti Jain*, V Manushree*

W3. NeurIPS 2021: XCI-Sketch

Aug 2021

[Oral] New in ML, **[Paper]** ML4CD, **[Paper]** CtrlGen, **[Poster]** DGM at **NeurIPS 2021**[GitHub](#) | [Paper](#)V Manushree, S Saxena, P Chowdhury, M Varma, H Rathod, Ankita Ghosh*, **Sahil Khose*****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](#)**Sahil Khose**, Abhiraj Tiwari, Ankita Ghosh**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](#)S Ramesh*, A Tiwari*, P Choubey*, S Kashyap*, **Sahil Khose***, K Lakara*, N Singh*, Ujjwal VermaSELECTED 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.

TEACHING EXPERIENCE

Graduate Teaching Assistant – CS 7647 Machine Learning with Limited Supervision ([course site](#))

Fall 2023

- Instructor: Prof. Judy Hoffman** | Guided 50 graduate students through state-of-the-art methods for visual learning with limited human supervision, mentoring 12 semester-long research projects from proposal to final evaluation.

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, LATALKS

SkyScenes: Synthetic-to-Real Generalization for Aerial Imagery, NASA S2A2 Annual Meeting, Georgia Tech

Jun 2023

An Efficient Modern Baseline for FloodNet VQA, New in ML @ ICML 2022

Jul 2022

Transformer-based Ensemble for Emotion Detection, WASSA @ ACL 2022

May 2022

XCI-Sketch: Extraction of Color Information from Images, New in ML @ NeurIPS 2021

Dec 2021

Semi-Supervised Classification & Segmentation on High-Res Aerial Images, CCAI @ NeurIPS 2021

Dec 2021

BERT Transformers for Extracting Health Information from Social Media, SMM4H @ NAACL 2021

Jun 2021