

# Sahil Khose

[sahilkhose.github.io](https://github.com/sahilkhose) ◊ [sahilkhose18@gmail.com](mailto: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

<b>University of California, Irvine, USA</b> <i>Ph.D. in Computer Science</i> <i>Advisor: Prof. Judy Hoffman</i>	2026 – Present GPA: 4.0/4.0
<b>Georgia Institute of Technology, Atlanta, USA</b> <i>Ph.D. in Computer Science</i> <i>Advisor: Prof. Judy Hoffman</i> <i>Transferred to University of California, Irvine (advisor relocation)</i>	2024 – 2026 GPA: 4.0/4.0
<b>Georgia Institute of Technology, Atlanta, USA</b> <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. Zsolt Kira, and Prof. Humphrey Shi</i>	2022 – 2024 GPA: 4.0/4.0
<b>Manipal Institute of Technology, Manipal, India</b> <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

<b>Georgia Institute of Technology, Atlanta, USA</b> <i>Graduate Research Assistant at Hoffman AI Lab</i>	Jan 2023 – Present <i>Advisor – Prof. Judy Hoffman</i>
<ul style="list-style-type: none"><li><b>WFMs for robotic manipulation:</b> Developing VLM-based benchmarks for automatic evaluation of world models. [P2]</li><li><b>OVSF-VLM:</b> 3D Scene Graph Generation model for spatial and semantic reasoning in real-world robotics tasks. [W8]</li><li><b>Multimodality:</b> Compact vision-audio LLM with MLP projectors and joint training, reduces cross-modal interference. [P1]</li><li><b>Syn-to-real:</b> Developed syn-to-real adaptation to raise off-road semantic segmentation performance.</li><li><b>SkyScenes:</b> Synthetic aerial benchmark that lifts model performance when transferring from syn-to-real. [C4] ECCV '24</li></ul>	
<b>Georgia Institute of Technology, Atlanta, USA</b> <i>Graduate Research Assistant at Neural Data Science Lab (NerDS)</i>	Spring 2023 <i>Advisor – Prof. Eva Dyer</i>
<ul style="list-style-type: none"><li><b>LatentDR:</b> a plug-and-play module to counter <b>diversity shift</b> without architecture changes. [C3] WACV '24</li></ul>	
<b>Indian Institute of Science, Bangalore, India</b> <i>AI Research Assistant at Artificial Intelligence and Robotics Lab</i>	Jul 2021 – Jul 2022 <i>Advisors – Prof. S. Sundaram &amp; Dr. Chandan Gautam</i>
<ul style="list-style-type: none"><li><b>Bachelor's Thesis:</b> Jointly addressing <b>domain shift</b> + <b>semantic shift</b> to recognize unseen classes in unseen domains.</li></ul>	
<b>Manipal Institute of Technology, Manipal, India</b> <i>Medical AI Research Assistant</i>	Apr 2021 – Oct 2022 <i>Advisor – Prof. Harish Kumar JR</i>
<ul style="list-style-type: none"><li>Developed a medical diagnosis system for <b>fovea segmentation</b> using semi-supervised segmentation. [C2]</li><li>Designed a <b>macular degeneration classification</b> system with interpretability for ophthalmology diagnosis. [C1]</li></ul>	
<b>Project MANAS – AI Robotics Research Team, MIT, Manipal, India</b> <i>AI Perception Developer</i>	Sep 2018 – May 2021 <i>GitLab   Website</i>
<ul style="list-style-type: none"><li><b>World Rank 1</b> at <b>IGVC 2019</b> (UGV) &amp; winner of the <b>Mahindra \$1M Challenge</b> (out of 153 self-driving car teams).</li><li>Implemented <b>Lane Detection, Speed Bump Detection, Driving Imitation System, Depth Map Generation.</b></li></ul>	

## ACHIEVEMENTS

### Research Awards

- 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

## Competitions

- **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 **Mahindra \$1Million Challenge (top 13 out of 153 teams in India)**. 2019

## ONGOING (PREPRINT)

---

### P2. VLMs for Robust Evaluation of World Models in Robotic Manipulation

*Under review*

**Sahil Khose**, Mengqi Zhang, 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.*

### 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

---

### W8. **CVPR 2026:** SPOT: Structured Prompting with Object-centric Tokens for open-world scene graphs

**[Poster]** *Multimodal Spatial Intelligence at CVPR 2026*

[Paper](#)

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

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

### 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

<b>W4. NeurIPS 2021: A Studious Approach to Semi-Supervised Learning</b> [Poster] ICBINB at <b>NeurIPS 2021</b> <b>Sahil Khose*</b> , Shruti Jain*, V Manushree*	Sep 2021 <a href="#">GitHub</a>   <a href="#">Paper</a>
<b>W3. NeurIPS 2021: XCI-Sketch</b> [Oral] New in ML, [Paper] ML4CD, [Paper] CtrlGen, [Poster] DGM at <b>NeurIPS 2021</b> V Manushree, S Saxena, P Chowdhury, M Varma, H Rathod, Ankita Ghosh*, <b>Sahil Khose*</b>	Aug 2021 <a href="#">GitHub</a>   <a href="#">Paper</a>
<b>W2. NeurIPS 2021: Semi-Supervised Classification &amp; Segmentation on High Resolution Aerial Images</b> [Spotlight Paper] Tackling Climate Change with ML at <b>NeurIPS 2021</b> <b>Sahil Khose</b> , Abhiraj Tiwari, Ankita Ghosh	May 2021 <a href="#">GitHub</a>   <a href="#">Paper</a>
<b>W1. NAACL 2021: BERT Transformers in Extraction of Health Information from Social Media</b> [Top Performer Award] Published in proceedings of <b>NAACL 2021</b> at SMM4H workshop S Ramesh*, A Tiwari*, P Choubey*, S Kashyap*, <b>Sahil Khose*</b> , K Lakara*, N Singh*, Ujjwal Verma	Apr 2021 <a href="#">GitHub</a>   <a href="#">Paper</a>

## SELECTED PROJECTS

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

## TEACHING EXPERIENCE

<b>Graduate Teaching Assistant – CS 7647 Machine Learning with Limited Supervision</b> <a href="#">Course site</a>	Fall 2023
<ul style="list-style-type: none"> <li><b>Instructor: Prof. Judy Hoffman</b>   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.</li> </ul>	

## PROFESSIONAL SERVICE

Conference Reviewer: **CVPR [2026, 2025]**, **NeurIPS [2026, 2025]**, **ECCV [2026, 2024]**  
 Workshop Reviewer: **NeurIPS-W 2025** (MATH-AI), **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

## TALKS

<b>SkyScenes: Synthetic-to-Real Generalization for Aerial Imagery</b> , NASA S2A2 Annual Meeting, Georgia Tech	Jun 2023
<b>An Efficient Modern Baseline for FloodNet VQA</b> , New in ML @ ICML 2022	Jul 2022
<b>Transformer-based Ensemble for Emotion Detection</b> , WASSA @ ACL 2022	May 2022
<b>XCI-Sketch: Extraction of Color Information from Images</b> , New in ML @ NeurIPS 2021	Dec 2021
<b>Semi-Supervised Classification &amp; Segmentation on High-Res Aerial Images</b> , CCAI @ NeurIPS 2021	Dec 2021
<b>BERT Transformers for Extracting Health Information from Social Media</b> , SMM4H @ NAACL 2021	Jun 2021