

Saksham Singh Kushwaha

Research Area : *Computer Vision, (Spatial) Audio, Multimodal*

Website: <https://sakshamsingh1.github.io/>

Email : sxk230060@utdallas.edu

EDUCATION

- **The University of Texas at Dallas** May 2027 (expected)
Ph.D. - Computer Science (Advisor: Dr. Yapeng Tian) GPA: 4.0/4.0
- **New York University, Courant** May 2023
M.S. - Computer Science GPA: 3.95/4.0
- **Indian Institute of Technology, Delhi** May 2018
B.Tech - Mathematics and Computing GPA: 7.97/10.0

EXPERIENCE

- **Adobe | PhD Research Intern** May 2025 - Aug 2025
Worked in content intelligence team for video editing Bangalore, India
- **Dolby Laboratories | PhD Research Intern** May 2024 - Aug 2024
Worked in Advanced Technology Group (ATG) to develop approaches for spatial audio generation SF, USA
- **Nvidia | Deep Learning Intern** May 2022 - Aug 2022
Worked in Product Security team to efficiently detect anomalous user-behavior in AWS accounts Remote, USA
- **Zomato | Machine Learning Engineer II** July 2018 - April 2021
Part of Search, User personalization, and Logistics teams Gurgaon, India

PUBLICATIONS

- **AVRobustBench: Benchmarking the Robustness of Audio-Visual Recognition Models at Test-Time** (paper)
S. Maharana, S. S. Kushwaha, B. Zhang, A. Rodriguez, S. Wei, Y. Tian, Y. Guo [NeurIPS D&B 2025]
- **VinTAGE: Joint Video and Text Conditioning for Holistic Audio Generation** (paper), (project page), (demo)
Saksham S. Kushwaha, Yapeng Tian [CVPR 2025]
- **Diff-SAGE: End-to-end spatial audio generation using diffusion models** (paper), (project page), (demo)
Saksham S. Kushwaha, Jianbo Ma, Mark Thomas, Yapeng Tian, Avery Bruni [ICASSP 2025]
- **Audio-Visual Dataset Distillation** (paper), (slides), (video)
Saksham S. Kushwaha, Siva Vasireddy, Kai Wang, Yapeng Tian [TMLR 2024]
- **A multimodal prototypical approach for unsupervised sound classification** (paper)
Saksham S. Kushwaha, Magdalena Fuentes [INTERSPEECH 2023]
- **Sound source distance estimation in diverse and dynamic acoustic conditions** (paper)
Saksham S. Kushwaha, Iran Roman, Magdalena Fuentes, Juan Pablo Bello [WASPAA 2023]
- **Analyzing the effect of equal-angle spatial discretization on sound event localization & detection** (paper)
Saksham S. Kushwaha, Iran R. Roman, Juan Pablo Bello [DCASE 2022]

ACADEMIC EXPERIENCE

- **Invited talks:** DL-MIR Workshop, Stanford University (slides)
- **Reviewer:** MLSP 2023, AAAI-24, ICASSP-24, NIPS-24, ACMMM-24, ICLR-25, ICML-25
- **Research Assistant:** Prof. Magdalena Fuentes (multimodal deep learning) [Aug'22-May'23], Prof. Raveesh Mayya (ML for digital policy change) [Aug'21-July'22], Prof. Yapeng Tian [Aug'24-Present]
- **Teaching Assistant:** Intro to python programming (Prof. Junpei Komiyama) [Jan'22-May'22], Discrete Mathematics (Prof. Simeon Ntafos) [May'23-Aug'23], Computer Science I (Prof. Scott Dollinger) [May'23-Aug'23]