

# Sagnik Majumder

✉ [sagnik@cs.utexas.edu](mailto:sagnik@cs.utexas.edu) | 🌐 [Webpage](#) | 🐙 [GitHub](#)

## EDUCATION

---

- DECEMBER 2025    **University of Texas at Austin (UT)-Austin**, Texas, United States  
Doctor of Philosophy and Master of Science in COMPUTER SCIENCE; GPA: 3.96/4.0
- JULY 2018    **Birla Institute of Technology and Science (BITS)-Pilani**, Pilani, Rajasthan, India  
Bachelor of Engineering (Hons.) in ELECTRONICS AND INSTRUMENTATION  
Thesis: "Neural Architecture Meta-learning via Reinforcement" | Advisor: [Prof. V. RAMESH](#)  
GPA: 9.55/10, Distinction and ranked 2<sup>nd</sup> out of 100 students

## INTERNSHIPS AND RESEARCH

---

- JUNE 2025 - present    **Research intern at Apple**  
Manager - [Dr. Ian Fasel](#)
- JAN 2020 - present    **Research assistant at UT Austin**  
Advisor - [Prof. Kristen Grauman](#)
- DEC 2022 - DEC 2024    **Visiting researcher at Meta AI**  
Manager - [Prof. Kristen Grauman](#)
- JUNE 2022 - DEC 2022    **Research scientist intern at Reality Labs Research, Meta**  
Manager - [Dr. Vamsi Krishna Ithapu](#)
- JAN 2018 - MAY 2019    **Research assistant at Goethe University**  
Advisor - [Prof. Visvanathan Ramesh](#)
- MAY 2017 - May 2019    **Research intern at Frankfurt Institute for Advanced Studies (FIAS)**  
Advisor - [Prof. Christoph Malsburg](#)

## SELECT PUBLICATIONS AND PREPRINTS

---

- Priyanka Mandikal, Jiaheng Hu, Shivin Dass, **Sagnik Majumder**, Roberto Martin Martin\*, Kristen Grauman\*. "Smash and Spread! Teaching Robots to Transform Objects via Spatial Progress". In submission. [\[Preprint\]](#), [\[Project Page\]](#).
- **Sagnik Majumder**, Tushar Nagarajan, Ziad Al-Halah, Kristen Grauman. "Switch-a-View: View Selection Learned from Unlabeled In-the-wild Videos". ICCV 2025. [\[arXiv\]](#), [\[Project Page\]](#).
- **Sagnik Majumder**, Tushar Nagarajan, Ziad Al-Halah, Reina Pradhan, Kristen Grauman. "Which Viewpoint Shows it Best? Language for Weakly Supervising View Selection in Multi-view Instructional Videos". CVPR 2025 (Highlight). [\[arXiv\]](#), [\[Project Page\]](#).
- Arjun Somayazulu, **Sagnik Majumder**, Changan Chen, Kristen Grauman. "ActiveRIR: Active Audio-Visual Exploration for Acoustic Environment Modeling". IROS 2024 (Oral). [\[arXiv\]](#), [\[Project Page\]](#).
- **Sagnik Majumder**, Ziad Al-Halah, Kristen Grauman. "Learning Spatial Features from Audio-Visual Correspondence in Egocentric Videos". CVPR 2024. [\[arXiv\]](#), [\[Project Page\]](#).
- Kristen Grauman, ..., **Sagnik Majumder**, ..., Michael Wray. "Ego-Exo4D: Understanding Skilled Human Activity from First- and Third-Person Perspectives". CVPR 2024 (Oral). [\[arXiv\]](#), [\[Project Page\]](#).

- **Sagnik Majumder**, Hao Jiang, Pierre Moulon, Ethan Henderson, Paul Calamia, Kristen Grauman\*, Vamsi Krishna Ithapu\*. “Chat2Map: Efficient Scene Mapping from Multi-Ego Conversations”. **CVPR 2023**. [\[arXiv\]](#), [\[Project Page\]](#).
- **Sagnik Majumder**, Changan Chen\*, Ziad Al-Halah\*, Kristen Grauman. “Few-Shot Audio-Visual Learning of Environment Acoustics”. **NeurIPS 2022**. [\[arXiv\]](#), [\[Project Page\]](#).
- **Sagnik Majumder**, Kristen Grauman. “Active Audio-Visual Separation of Dynamic Sound Sources”. **ECCV 2022**. [\[arXiv\]](#), [\[Project Page\]](#).
- **Sagnik Majumder**, Ziad Al-Halah, Kristen Grauman. “Move2Hear: Active Audio-Visual Source Separation”. **ICCV 2021**. [\[arXiv\]](#), [\[Project Page\]](#).
- Changan Chen, **Sagnik Majumder**, Ziad Al-Halah, Ruohan Gao, Santhosh K. Ramakrishnan, Kristen Grauman. “Learning to Set Waypoints for Audio-Visual Navigation”. **ICLR 2021**. [\[Publication\]](#), [\[Project Page\]](#).
- **Sagnik Majumder**, Chinmoy Samant, Greg Durrett. “Model Agnostic Answer Reranking System for Adversarial Question Answering”. **EACL 2021 Student Research Workshop**. [\[Publication\]](#).

## INVITED TALKS AND PRESENTATIONS

---

<a href="#">EgoVis at CVPR 2025</a>	Which Viewpoint Shows it Best? Language for Weakly Supervising View Selection in Multi-view Instructional Videos
<a href="#">EgoVis at CVPR 2024</a>	Learning Spatial Features from Audio-Visual Correspondence in Egocentric Videos
<a href="#">Sight and Sound at CVPR 2024</a>	Learning Spatial Features from Audio-Visual Correspondence in Egocentric Videos
<a href="#">Sight and Sound at CVPR 2023</a>	Chat2Map: Efficient Scene Mapping from Multi-Ego Conversations
<a href="#">NSA Lab at JHU</a>	Efficiently understanding 3D scenes using sight and sound
<a href="#">ECCV AV4D workshop 2022</a>	Active Audio-Visual Separation of Dynamic Sound Sources
<a href="#">Sight and Sound at CVPR 2022</a>	Active Audio-Visual Separation of Dynamic Sound Sources
<a href="#">Embodied AI seminar at Meta AI</a>	Active Audio-Visual Separation of Dynamic Sound Sources

## PROFESSIONAL SERVICE

---

Workshop co-organizer: CVPR Embodied AI workshop 2022-23  
Reviewer: CVPR; ECCV; ICCV; NeurIPS; ICML; AAAI; RA-L; ICRA; BMVC; TNNLS

## ACADEMIC HONORS AND ACHIEVEMENTS

---

JULY 2025	Received the Professional Development Award from UT Austin for presenting my research at CVPR 2025
MAY 2025	Recognized as Outstanding Reviewer (top 5.6% of 12,593 reviewers) at CVPR 2025
JULY 2024	Received the Professional Development Award from UT Austin for presenting my research at CVPR 2024
JULY 2023	Received the Professional Development Award from UT Austin for presenting my research at CVPR 2023
NOVEMBER 2020	TOEFL iBT: 118 (READING: 29, LISTENING: 30 SPEAKING: 29, WRITING: 30)
JULY 2018	GRE: 334 (QUANTITATIVE: 170, VERBAL: 164, AWA: 5.0)
JAN 2016 - JUN 2018	Received academic excellence scholarship from BITS Pilani for 5 consecutive semesters
AUGUST 2017	Offered full-time role as machine learning engineer (software development) by Goldman Sachs India
MARCH 2017	Secured 2 <sup>nd</sup> place in paper presentation at APOGEE, BITS Pilani technical festival
DECEMBER 2016	Received DAAD WISE scholarship 2017 for research internship in Germany
JUNE 2014	Ranked in top 0.50% in IIT-JEE and 64 in WBJEE
FEBRUARY 2014	Offered KVPY fellowship by the Govt. of India

## COURSEWORK

---

- Graduate: **Deep Learning Seminar; Reinforcement Learning: Theory & Practice; Robot Learning; Natural Language Processing; Spoken Language Technologies; Math in Deep Learning; Statistical Models for Health and Behavioral Sciences;**  
Algorithms: Techniques and Theory; Programming Languages;
- Undergraduate: Neural Networks & Fuzzy Logic; Machine Learning; Advanced Calculus; Linear Algebra and Complex Variables; Probability and Statistics; Computer Programming; Operating Systems; Object Oriented Programming; Advanced Computer Architecture; Algorithms and Complexity; Data Structures; Discrete Mathematics
- MOOC: [Stanford's CS231n: Convolutional Neural Networks for Visual Recognition](#);  
[Stanford's CS224n: Natural Language Processing with Deep Learning](#);  
[UC Berkeley's CS294: Deep Reinforcement Learning](#)

## SOFTWARE SKILLS

---

- Programming Language: Python; C; C++; Java; Matlab
- Autodifferentiation Framework: PyTorch; Tensorflow; Caffe
- Python Package: Numpy; Scipy; SK-learn; Matplotlib; Seaborn; Plotly
- Operating System: Linux (Debian, Ubuntu); MS Windows
- Distributed Version Control: Git
- Document Preparation:  $\text{\LaTeX}$ ; MS Word

## TEACHING EXPERIENCE

---

- SPRING 2025: Teaching assistant for "CS 381V: Visual Recognition" at UT Austin
- SEMESTER 1, 2017-18: Teaching assistant for "Neural Networks and Fuzzy Logic" at BITS Pilani