

# Kumar Ashutosh

✉ kumar.ashutosh@utexas.edu

🌐 <https://ashutoshkr.me>

## Education

2021 – Present	📖 <b>The University of Texas at Austin</b> Ph.D. in Computer Science Advisor: Kristen Grauman Research area: Computer Vision, Machine Learning	Austin, TX
2016 – 2021	📖 <b>Indian Institute of Technology Bombay</b> Master's in Electrical Engineering Thesis title: <i>3D Shape Reconstruction with View-Planning</i> .	Mumbai, India
	📖 <b>Indian Institute of Technology Bombay</b> Bachelor's in Electrical Engineering, Minor in Computer Science	

## Work Experience

Present	📖 <b>UT Austin</b> , Graduate Research Assistant	Austin, TX
Summer 2025	📖 <b>Gen AI, Meta</b> , Research Scientist Intern	New York, NY
2023-25	📖 <b>FAIR, Meta</b> , Visiting Researcher	Austin, TX
Summer 2022	📖 <b>FAIR, Meta</b> , Research Intern	New York, NY
Winter 2019	📖 <b>360World</b> , AR/VR Developer Intern	Budapest, Hungary
Summer 2019	📖 <b>Sony Corporation</b> , Research Engineer Intern	Kanagawa, Japan
Summer 2018	📖 <b>National University of Singapore</b> , Research Intern	Singapore
	📖 <b>Google Summer of Code</b> , Developer	Remote

## Research Publications

### Preprints

- 1 **K. Ashutosh**, Y. Gandelsman, X. Chen, I. Misra, and R. Girdhar, "LLMs can see and hear without any training," 2025. arXiv: 2501.18096 [cs.CV].
- 2 **K. Ashutosh**, R. Girdhar, L. Torresani, and K. Grauman, "What You Say Is What You Show: Visual Narration Detection in Instructional Videos," 2023. arXiv: 2301.02307 [cs.CV].

### Conference Proceedings

- 1 **K. Ashutosh**, T. Nagarajan, G. Pavlakos, K. Kitani, and K. Grauman, "ExpertAF: Expert Actionable Feedback from Video," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2025, [🔗](#).
- 2 **K. Ashutosh**, G. Pavlakos, and K. Grauman, "Flction: 4D Future Interaction Prediction from Video," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2025, **Highlight paper (Top 3%)** [🔗](#).
- 3 **K. Ashutosh**, Z. Xue, T. Nagarajan, and K. Grauman, "Detours for Navigating Instructional Videos," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2024, **Highlight paper (Top 2.8%)** [🔗](#).
- 4 C. Chen, **K. Ashutosh**, R. Girdhar, D. Harwath, and K. Grauman, "Discovering sounding actions in video with multimodal consensus," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2024.

- 5 K. Grauman, A. Westbury, L. Torresani, *et al.*, “Ego-Exo4D: Understanding Skilled Human Activity from First- and Third-Person Perspectives,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2024, **Oral paper (Top 0.8%)** [🔗](#).
- 6 Z. Xue, **K. Ashutosh**, and K. Grauman, “Learning Object State Changes in Videos: An Open-World Perspective,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2024, [🔗](#).
- 7 **K. Ashutosh**, S. Ramakrishnan, T. Afouras, and K. Grauman, “Video-mined task graphs for keystep recognition in instructional videos,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2023, [🔗](#).
- 8 **K. Ashutosh**, R. Girdhar, L. Torresani, and K. Grauman, “HierVL: Learning Hierarchical Video-Language Embeddings,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2023, **Highlight paper (Top 2.5%)** [🔗](#).
- 9 A. Jaiswal, **K. Ashutosh**, J. F. Rousseau, Y. Peng, Z. Wang, and Y. Ding, “RoS-KD: A Robust Stochastic Knowledge Distillation Approach for Noisy Medical Imaging,” in *2022 IEEE International Conference on Data Mining (ICDM)*, Dec. 2022, [🔗](#).
- 10 **K. Ashutosh**, S. Kumar, and S. Chaudhuri, “3D-NVS: A 3D Supervision Approach for Next View Selection,” in *2022 26th International Conference on Pattern Recognition (ICPR)*, Aug. 2022, [🔗](#).
- 11 **K. Ashutosh**, J. Nair, A. Kagrecha, and K. Jagannathan, “Bandit algorithms: Letting go of logarithmic regret for statistical robustness,” in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2021, **Oral presentation (Top 3%)** [🔗](#).
- 12 **K. Ashutosh**, “Hardware performance analysis of mobile-based augmented reality systems,” in *2020 International Conference on Computational Performance Evaluation (ComPE)*, 2020, [🔗](#).
- 13 **K. Ashutosh**, S. Consul, B. Dedhia, P. Khirwadkar, S. Shah, and S. Kalyanakrishnan, “Lower bounds for policy iteration on multi-action mdps,” in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020, [🔗](#).
- 14 R. Bose, **K. Ashutosh**, J. Li, A. Dragomir, N. Thakor, and A. Bezerianos, “A multilayer network approach for studying creative ideation from eeg,” in *Brain Informatics*, Springer International Publishing, 2018, [🔗](#).

## Professional Service

---

### Reviewer


CVPR 2024	📌 <b>Outstanding reviewer (Top 2% out of 10k reviewers)</b>
NeurIPS 2024	📌 <b>Top reviewer (Top 8% out of 14k reviewers)</b>
2023-24	📌 The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)
2023	📌 The IEEE/CVF International Conference on Computer Vision (ICCV)
2022-24	📌 European Conference on Computer Vision (ECCV)
2024	📌 Neural Information Processing Systems (NeurIPS)
2024-25, 22	📌 The IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)
2024	📌 Asian Conference on Computer Vision (ACCV)
2025	📌 AAAI Conference on Artificial Intelligence (AAAI)

### Teaching Assistant



Spring 2025	📌 <i>Visual Recognition</i> in Computer Science, UT Austin
Spring 2021	📌 <i>Matrix Computations</i> in Electrical Engineering, IIT Bombay

## Professional Service (continued)







---

Autumn 2020     *Applied Linear Algebra* in Electrical Engineering, IIT Bombay

### Responsibilities









- 2019 – 2020     Department Academic Mentor Coordinator in Electrical Engineering, IIT Bombay
- 2018 – 2021     Institute and Department Academic Mentor, IIT Bombay
- 2017 – 2018     Convener of *Web n Coding Club*, IIT Bombay

### Open Source

- 2023     Contributed to PyTorch codebase 
-  Open sourced the implementation of our CVPR 2023 paper (HierVL): 
- 2017     Contributed 2k+ lines of code to scikit-learn, a popular ML package 







## Awards and Achievements

---

- 2023     Professional Development Award by UT Austin to attend NeurIPS 2023
-  Professional Development Award by UT Austin to attend CVPR 2023
- 2020     Department Color by IIT Bombay for valuable contribution to the mentorship program
- 2017     Invited to the Republic Day Parade as a guest of the Hon'ble Prime Minister of India
- 2016     Rashtrapati Puraskar (President's Award) by the Hon'ble President of India for Scouting
-  99.97 (out of 1.2M) and 99.14 percentile (out of 0.15M) in JEE Mains and Advanced
- 2015     Qualified Indian National Mathematical Olympiad and attended selection camp for IMO
-  Cleared NTSE and KVPY scholarship exams organized by the Govt. of India








## Press Coverage

---

- 2023     Meta AI's coverage of our CVPR 2023 paper : , , , 




## Talks

---

- 2024     Invited talk at IIT Delhi in Jan titled “Long-video understanding with text supervision”.
- 2023     Highlight presentation of our paper at CVPR 2023 
-  Invited talk at International Workshop on Large Scale Holistic Video Understanding, CVPR 2023
- 2021     Oral presentation of our work at AISTATS 2021 
- 2020     Invited Talk on Augmented Reality Applications at Electronics and Robotics Club, IIT Bombay




## Technical Skills

---

- Languages     Python, C++, C, HTML, CSS,  $\text{\LaTeX}$
- ML Tools     Huggingface, Deepspeed, Fairseq, SLURM, PyTorch, Tensorflow, Git
- Development     React, Jekyll, Android Studio, Xcode, Unity

# Extracurricular activities

---

- Sports      Football (Soccer), Squash, Cricket, Tennis, Badminton
- Activities      Hiking, Running
- Music      Guitar, Piano/Keyboard