Arka Sadhu

http://theshadow29.github.io • ark.sadhu2904@gmail.com • LinkedIn • Github • GScholar • Long CV
Meta, Sunnyvale, CA, 94086

Work Experience

- **Meta, Sunnyvale, CA, USA** (June'24 - Present)
  Applied Research Scientist in Surreal Team Working in surreal team on next generation glasses for ego-centric videos to power contextual AI agents. Contributed to multiple projects from Anonymization, LLM Agents, Retrieval Augmentation Generation (RAG), Context Compression, AI for glasses. Our project was shown to all top-executives in Meta including Mark Zuckerberg, Andrew Bosworth, Chris Cox gaining high praise. My contribution was particularly towards the LLM Agents.

Education

PhD Computer Science University of Southern California, Los Angeles, CA, USA Aug'18-May'24
- **Advisor: Prof. Ram Nevatia • CGPA:** 4.0/4.0
- **Annenberg Fellow 2018-22**

Bachelor of Electrical Engineering Indian Institute of Technology Bombay, India Jun'14-May'18
- **Honors in Electrical Engineering and Minor in Computer Science • CGPA:** 9.44/10

High School R.N. Podar School CBSE, Santacruz, Mumbai Apr'12-Mar'14

Selected Peer-Reviewed Publications

- DiVE-k: Differential Visual Reasoning for Fine-grained Image Recognition [Paper]
  Raja Kumar, **Arka Sadhu**, Ram Nevatia **Under Submission ICLR 2026**
- Leveraging Task-Specific Pre-Training to Reason across Images and Videos [Paper]
  Arka Sadhu, Ram Nevatia **WACV 2024**
- Gradient-based Editing of Memory Examples for Online Task-free Continual Learning [Paper][Code]
  Xisen Jin, **Arka Sadhu**, Junyi Du, Xiang Ren **Neurips 2021**
- Visual Semantic Role Labeling for Video Understanding [Paper][Code]
  Arka Sadhu, Tanmay Gupta, Mark Yatskar, Ram Nevatia, Aniruddha Kembhavi **CVPR 2021**
- Video Question Answering with Phrases via Semantic Roles [Paper][Code]
  Arka Sadhu, Kan Chen, Ram Nevatia **NAACL 2021**
- Video Object Grounding using Semantic Roles in Language Description [Paper] [Code]
  Arka Sadhu, Kan Chen, Ram Nevatia **CVPR 2020**
- Zero-Shot Grounding of Objects from Natural Language Queries. [Paper] [Code]
  Arka Sadhu, Kan Chen, Ram Nevatia **ICCV 2019 (Oral)**

Academic Experiences and Services

Reviewer: CVPR'21,22,23,24,25,26, **ECCV'22,24**,
ICCV'21,23,25,**ACL'21,22,23,24,25**,**EMNLP'21,22,23,24,25**, **WACV'20,21,22,23,24,25,26**, **IROS'21**,
ICPR'21,22, **BMVC'20,21,22,23,24,25**, **AAAI'21,22,26**, **ICLR'22,24,26**, **ICML'23,25**,
Neurips'22,23,24,25, **ARR 21,22,23,24,25**, **ICRA'26**, **AURO**, **TPAMI**, **IJCV**
Keynote Speaker: Workshop on Media Forensics (CVPR'19)
Workshop Organizer: Workshop on Synthetic Realities Audio Visual Fakes (ICML'19)
TA: Advanced Computer Vision (CSCI 677) Fall'19, Analysis of Algorithms (CSCI 570) Spring'21
HackwithBay Financial Agents Hackathon, Finalist Judge (Oct'25) [Website]

Notable Projects

- **VidSitu Leaderboard**[Website] Leaderboard for VidSitu test sets.
- **Research Advice List**[Github]: A curated list of presentations and write-ups on research and paper writing.
- **Awesome-Grounding** [Github]: Curated list of papers on grounding language in vision.
- **Visual Commonsense using BERT in Pytorch** [Github]: improved BERT baselines on VCR dataset.
- **Inference Networks for SPENs** [Report] [Github]: Tensorflow implementation of ICLR'18 paper Learning Approximate Inference Networks for Structured Prediction.
- **Voice Conversion Using GANs** [Report] [Github]: Tensorflow implementation of Interspeech'17 paper with additional sentence encoding scheme.

Academic and Technical Achievements

- Awarded Annenberg grant for presenting in Annenberg Symposium.
- Placed 3rd at the IFood Challenge which was a part of the FGVC5 at CVPR 2018. [Code]
- Awarded Annenberg Fellowship (2018), Viterbi-India Scholarship (2017), Aalto fellowship (2016).
- Secured All India Rank 134 in JEE Advance 2014 among 1.5 lakh selected candidates.

Skills

- **Programming Languages:** Proficient in: C++, Python, Matlab, Bash, Embedded-C, Assembly
- **Deep Learning Frameworks:** Pytorch, Tensorflow, Caffe.
- **Computer Vision Frameworks:** OpenCV, Scikit-Image, Scikit-Learn
- **NLP Frameworks:** Spacy, StanfordCoreNLP