

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
- **Meta AI, Menlo Park, USA** (May'22 - Aug'22)
Research Intern with Animesh Sinha, Licheng Yu, Hugo Chen and Ning Zhang
- **Allen Institute for Artificial Intelligence (AI2), Seattle, USA** (May'21 - Aug'21)
Research Intern with Tanmay Gupta, and Aniruddha Kembhavi
- **Allen Institute for Artificial Intelligence (AI2), Seattle, USA** (May'20 - Aug'20)
Research Intern with Mark Yatskar, Tanmay Gupta, and Aniruddha Kembhavi
- **Wadhvani Institute of AI, Mumbai, India** (May'18 - Aug'18)
Research Intern with Prof. Subhasis Chaudhuri and Rahul Panicker

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**

SELECTED FIRST-AUTHOR PEER-REVIEWED PUBLICATIONS

- Detect Anything: An Open-Vocabulary Object Detection Model with Weak Supervision
Arka Sadhu, Ram Nevatia **Under Submission**
- Leveraging Task-Specific Pre-Training to Reason across Images and Videos [Paper]
Arka Sadhu, Ram Nevatia **WACV'24**
- Unaligned Video-Text Pre-training using Iterative Alignment [Paper]
Arka Sadhu, Licheng Yu, Animesh Sinha, Yu Chen, Ram Nevatia, Ning Zhang
- 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

- Reviewer: CVPR'21,22,23,24, ECCV'22,24, ICCV'21,23, ACL'21,22,23,24, EMNLP'21,22,23, WACV'20,21,22,23,24, IROS'21, ICPR'21,22, BMVC'20,21,22,23, AAAI'21,22, ICLR'22, Neurips'22,23, ARR Nov'21-Nov'22, AURO, TPAMI, IJCV**
- 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**

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