

# MEDHINI G. NARASIMHAN

medhini95@gmail.com | (217) 904-6942 | <https://medhini.github.io/>

## EMPLOYMENT

---

**Staff Research Scientist, GenMedia, Google Deepmind** Aug 2023 - **Present**

- *Core work:* [Veo](#), latent diffusion video generation model.
- Veo 2 - Model scaling laws, pre-training, data mixtures, image-to-video, and post-training.
- Veo 3 - Evals (lead), audio conditioning, and joint video-audio post-training.
- Lead for training Veo 1-small and Veo 1.5 - efficient Veo models served in [YouTube DreamScreen](#), [VideoFX](#) and [Google Cloud Vertex AI](#).

**Research Intern, Meta AI Research, CA** May 2022 - Dec 2022

- *Project:* Learning visual representations for tasks in instructional videos  
*Mentors:* Dr. N Zhang, Prof. T Berg
- Trained VideoTaskformer, a large multimodal vision transformer which achieves best performance in action forecasting and task verification in videos

**Research Intern, Google Research, CA** May 2021 - May 2022

- *Project:* Generating visual summaries of long instructional videos  
*Mentors:* Dr. A. Nagrani, Prof. C. Schmid
- Developed a multimodal summarization model that extracts relevant scenes in long video clips using audio, text, and visual information.

## EDUCATION

---

**Doctor of Philosophy, Computer Science** Aug 2019 - Aug 2023

University of California at Berkeley

**GPA: 3.96/4.00**

*Advisor:* Prof. T. Darrell

*Thesis Title:* Multimodal Long-Term Video Understanding

**Master of Science, Computer Science**

University of Illinois at Urbana-Champaign

Aug 2017 - May 2019

**GPA: 4.00/4.00**

*Advisors:* Prof. S. Lazebnik & Prof. A. Schwing

*Thesis Title:* Visual question answering using external knowledge

**Bachelor of Technology, Information Technology**

National Institute of Technology, Karnataka, India

Jul 2013 - May 2017

**GPA: 9.50/10.00**

*Thesis Title:* Automatic Generation of Sentential

## PEER REVIEWED PUBLICATIONS

---

- S. Subramaniam, **M. Narasimhan**, A. Nagrani, A. Zeng, C. Schmid, D. Klein, T. Darrell  
“Modular Visual Question Answering via Code Generation”, ACL 2023.
- C. Shan, E. Tran, **M. Narasimhan**, S. Subramaniam, D. Klein, T. Darrell  
“LUSE: Using LLMs for Unsupervised Step Extraction in Instructional Videos”,  
**Spotlight** at CVEU@ICCV 2023.
- **M. Narasimhan**, L. Yu, S. Bell, N. Zhang, T. Darrell  
“Learning and Verification of Task Structure from Instructional Videos”, CVEU@ICCV 2023.
- **M. Narasimhan**, A. Nagrani, C. Sun, T. Darrell, A. Rohrbach, C. Schmid  
“TL;DW? Summarizing Instructional Videos with Task Relevance & Cross-Modal Saliency”, ECCV 2022.

- **M. Narasimhan**, A. Rohrbach, T. Darrell, “*CLIP-It! Language-Conditioned Video Summarization*”, NeurIPS 2021.
- **M. Narasimhan**, S. Ginosar, A. Owens, A. Efros, T. Darrell, “*Strumming to the Beat: Audio-Conditioned Contrastive Video Textures*”, WACV 2022.
- **M. Narasimhan**, E. Wijmans, X. Chen, T. Darrell, D. Batra, D. Parikh, A. Singh, “*Seeing the Un-Scene: Learning Amodal Semantic Belief Maps for Room Navigation*”, ECCV 2020.
- **M. Narasimhan**, S. Lazebnik, and A. Schwing, “*Out of the Box: Reasoning with Graph Convolution Nets for Factual Visual Question Answering*”, NeurIPS 2018.
- **M. Narasimhan**, and A. Schwing, “*Straight to the Facts: Learning Knowledge Base Retrieval for Factual Visual Question Answering*”, ECCV 2018.
- **M. Narasimhan**, B. Balasubramanian, S. Kumar, and N. Patil, “*EGA-FMC: Enhanced Genetic Algorithm based Fuzzy K-Modes Clustering for Categorical Data*”, International Journal of Bio-Inspired Computation 2018.
- **M. Narasimhan**, and S. Kamath, “*Dynamic video anomaly detection and localization using sparse denoising autoencoders*”, Multimedia Tools and Applications 2017.
- **M. Narasimhan**, G. Vietri, A. Mehta, F. Rajabli, V. Aguiar-Pulido, K. Mathee, and G. Narasimhan, “*Predicting Symptom Severity and Contagiousness of Respiratory Viral Infections*”, ISCB 2016.

---

## SKILLS

**Areas:** Computer Vision, Natural Language Processing, Deep Learning, Machine Learning

**Languages and Frameworks:** Python, PyTorch, Tensorflow, JAX

---

## AWARDS AND ACHIEVEMENTS

- Co-founded the [Berkeley AI Climate Research Initiative](#).
- Best Student Paper Award at Winter Applications of Computer Vision (WACV) 2022.
- Google AI for Social Good Program, 2021 - awarded funding for an AI for conservation project in collaboration with WCS Nigeria.
- Snap Research Scholar, 2019 - awarded to 8 students for outstanding research in Computer Vision.
- Siebel Scholar, Class of 2019 - awarded annually for academic excellence and demonstrated leadership to over 90 top students from the worlds leading graduate schools.
- AMD Best Undergraduate Thesis Award 2017 for best thesis across multiple schools.
- Huawei Merit Scholarship 2017 for class position and academic merit in NITK (*top 2 in class*).

---

## LEADERSHIP, VOLUNTEERING, & EXTRA CURRICULARS

- **Reviewer:** CVPR, ECCV, ICCV, NeurIPS, ICML, and TPAMI.
- Head of the Computing Society at ACM NITK branch, 2016-2017. Organized hackathons, tech events, talks, and mentoring programs for undergraduate students.
- Teaching Assistant for courses at UIUC and UC Berkeley: Computer Vision, Introduction to Computer Science, and Introduction to AI.
- *AI for Social Good:* Monitoring coral reef health by using computer vision video models to detect fish behavior.
- *AI for Social Good:* Animal census from camera trap images using unsupervised re-identification/recognition methods.
- *AI for Social Good:* Video to text for the blind - Developed a deep learning framework to automatically generate descriptions from videos as a part of the Microsoft Video-to-Text challenge.