

SHANMUKHA RAMAKRISHNA VEDANTAM

313 1st Street, Apt 4C, Hoboken, NJ, USA

+1-650-407-7845 • ramav@fb.com • <https://vrama91.github.io/>

CURRENT WORK

Research Scientist, Feb 2019 - Present

Facebook AI Research (FAIR), NY

RESEARCH INTERNSHIPS

Research Internship, Summer 2018

Microsoft Research Cambridge, United Kingdom

- Generative Models for Concept Learning
- Supervisor: Nate Kushman, Matthew Johnson, and Sebastian Nowozin, Microsoft Research

Research Internship, Summer 2017

Facebook AI Research (FAIR), Menlo Park, CA

- Supervisor: Devi Parikh, Facebook AI Research/Georgia Tech

Research Internship, Winter 2017

Machine Perception Group, Google Research, Mountain View, CA

- Grounded latent variable generative models for images and semantics.
- Supervisor: Kevin Murphy, Google Research

Research Internship, Summer 2016

Machine Perception Group, Google Research, Mountain View, CA

- Worked on a system to explain class discrimination conditioned on an image, using natural language
- Given an image, a target category and a distractor, explain why the image contains the target
- Supervisor: Gal Chechik and Samy Bengio, Google Research and Google Brain

Research Internship, Summer 2014

Center for Visual Computing, Ecole Centrale de Paris/ INRIA - Saclay, France

- Worked on Loopy Part Models for Face Detection
- Supervisor: Prof. Iasonas Kokkinos, Ecole Centrale de Paris

EDUCATION

- | | |
|---|-----------|
| • Georgia Institute of Technology, United States | 2017-2018 |
| Ph.D, Computer Science | |
| Advisor: Prof. Devi Parikh | |
| Thesis: Interpretation, Grounding and Imagination for Machine Intelligence | |
| • Virginia Polytechnic Institute and State University, United States | 2013-2016 |
| M.S., Computer Engineering | |
| Advisor: Prof. Devi Parikh | |
| Specialization: Computer Vision and Machine Learning | |
| • International Institute of Information Technology (IIIT), Hyderabad, India | 2009-2013 |
| Bachelor of Technology, Electronics and Communication Engg. | |
| Advisor: Prof. K. Madhava Krishna | |
| Specialization: Vision for Robotics | |

HONORS AND ACHIEVEMENTS

1. Awarded the Google PhD Fellowship in Machine Perception, Speech Technology and Computer Vision
[One out of 5 awardees selected across North America, Europe and the Middle East](#) 2018
2. Awarded the ICLR travel award for attending the International Conference on Learning Representations 2018
3. Finalist for the Adobe Research Fellowship 2018

4. Outstanding reviewer award at CVPR
Awarded to 130 reviewers in the CVPR reviewer pool 2017
5. Awarded travel grant of USD 1000 for CVPR, 2017 under Google's Archimedes program 2017
6. Finalist for the Adobe Research Fellowship 2016
7. Best Discussion Participant Award, Advanced Computer Vision Course, Virginia Tech Spring, 2014
8. Mentioned in Dean's List for excellence in academics at IIIT Hyderabad
Monsoon, 2009 & 2011, Spring 2012
9. Winner of Judges award and Peer award at Siemens CTT Intern Tech Challenge 2012
10. 3rd in global aerospace competition CANSAT organized by NASA, AAS and AIAA 2011
11. Top 20 rank in Regional Mathematics Olympiad Organized by National Board for Higher Mathematics (NBHM) from Gujarat State (qualified for Indian National Mathematics Olympiad) 2008
12. Finalist for the Bal Shree honor, conferred by the President of India for outstanding creativity in Science 2008
13. Awarded Chacha Nehru Scholarship for Artistic and Innovative Excellence from National Council of Educational Research and Training (NCERT) 2008
14. Attained All India Rank 134 in National Science Olympiad 2006
15. All India Rank 13 in Indian National Cartography Association (INCA) Map Quiz 2006

PUBLICATIONS

1. **Probabilistic Neural-Symbolic Models for Interpretable Visual Question Answering:** Ramakrishna Vedantam, Karan Desai, Stefan Lee, Marcus Rohrbach, Dhruv Batra, Devi Parikh. *International Conference on Machine Learning (ICML)*, 2019
2. **Generative Models of Visually Grounded Imagination:** Ramakrishna Vedantam, Ian Fischer, Jonathan Huang, Kevin Murphy. *International Conference on Learning Representations (ICLR)*, 2018 [Top 10%]
3. **Grad-CAM: Why did you say that? Visual Explanations from Deep Networks via Gradient-based Localization:** Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra. *International Conference on Computer Vision (ICCV)*, 2017
Also presented at *NIPS Workshop on Interpretable Machine Learning in Complex Systems*, 2016
4. **Sound-Word2Vec: Learning Word Representations Grounded in Sounds:** Ashwin K. Vijayakumar, Ramakrishna Vedantam, Devi Parikh. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2017
5. **Counting Everyday Objects in Everyday Scenes:** Prithvijit Chattopadhyay*, Ramakrishna Vedantam*, Ramprasaath R. Selvaraju, Dhruv Batra, Devi Parikh. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017 (**Spotlight**) [Top 8.2%]
6. **Context-aware Captions from Context-agnostic Supervision:** Ramakrishna Vedantam, Samy Bengio, Kevin Murphy, Devi Parikh, Gal Chechik. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017 (**Spotlight**) [Top 8.2%]
Also presented as an Oral at the *Bay Area Machine Learning Symposium (BayLearn)*, 2017.
7. **Adopting Abstract Images for Semantic Scene Understanding:** C. Lawrence Zitnick, Ramakrishna Vedantam, and Devi Parikh. *Special Issue on the best papers at the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2016

8. **Visual Word2Vec (vis-w2v): Learning Visually Grounded Word Embeddings using Abstract Scenes:** Satwik Kottur*, Ramakrishna Vedantam*, José Moura, and Devi Parikh. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016*
9. **Learning Common Sense through Visual Abstraction:** Ramakrishna Vedantam*, Xiao Lin*, Tanmay Batra, C. Lawrence Zitnick, and Devi Parikh. *IEEE International Conference on Computer Vision (ICCV), 2015*
Also presented as an oral at *1st Workshop on Object Understanding for Interaction*, colocated with *ICCV, 2015*
10. **CIDEr: Consensus-based Image Description Evaluation:** Ramakrishna Vedantam, C. Lawrence Zitnick, and Devi Parikh. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015*
* Equal Contribution

ArXiv Manuscripts

- **Microsoft COCO Captions: Data Collection and Evaluation Server:** Xinlei Chen, Hao Fang, Tsung-Yi Lin, Ramakrishna Vedantam, Saurabh Gupta, Piotr Dollar, C. Lawrence Zitnick. [April, 2015]

Teaching Assistant, Fall 2013 and Spring 2014

Bradley Department of ECE at Virginia Tech, United States

- Teaching assistant for ECE 2504, Introduction to Computer Engineering
- Responsible for grading, conducting office hours and project evaluation for a class of 60+ students

Internship, Summer 2012

Siemens AG - Corporate Research and Technologies, India

- Implemented a dense 3D reconstruction pipeline for aerial images using tracking, structure-from-motion and multi-view stereo software
- Supervisor: Dr. Antony Priyakumar, Siemens Corporate Research - Bangalore

Teaching Assistant, Monsoon 2011 and Spring 2013

International Institute of Information Technology, India

- Teaching Assistant for a freshman year course on Digital Logic and Processor design
- Teaching Assistant for Introduction to Humanities

PROFESSIONAL SERVICES

- **Conference:** Reviewer for ICVGIP 2014 and 2018, ICCV 2015-Present, CVPR 2016-Present, ECCV 2016-Present, ACCV 2016, ICVGIP 2016, BMVC 2017, NIPS 2017-Present, ICLR 2018, ICML 2018
- **Journals:** Reviewer for International Journal of Computer Vision (IJCV), IEEE Transactions on Image Processing, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), Neural Computation, Computer Speech and Language

ADVISING

- Ananya Raval (MS Student), Georgia Institute of Technology, Fall 2017
- Satwik Kottur (Intern), Carnegie Mellon University (co-advised with Devi Parikh) Summer 2015
- Prithvijit Chattopadhyay (Intern), Delhi Technical University (co-advised with Devi Parikh) Summer 2015

OPEN SOURCE CONTRIBUTIONS

- Developer on the coco-caption project which implements commonly used image captioning metrics such as CIDEr, METEOR, BLEU, and ROUGE-L.

- Developer of the CIDEr project which implements the two versions of CIDEr (CIDEr and CIDEr-D) from our CVPR'15 paper on Consensus Based Image Description Evaluation.

TALKS

Connecting Vision and Language via. Interpretation, Grounding, and Imagination

- University of Oxford (July, 2018)
- Google DeepMind, London (May, 2018)
- Microsoft Research, Cambridge (May, 2018)
- Facebook AI Research, Menlo Park (April, 2018)
- University of California, Berkeley (April, 2018)
- Google, Mountain View (May, 2018)
- Allen Institute for AI Research (AI2), Seattle (May, 2018)
- Toyota Technological Institute (TTI), Chicago (April, 2018)
- Indian Institute of Science (IISc), Bangalore (December, 2017)
- International Institute of Information Technology (IIIT), Hyderabad (December, 2017)

OTHER PROJECTS

- **Loopy Part Models for Face Detection** *INRIA-Saclay*
 Advisor: Iasonas Kokkinos and Dhruv Batra Summer 2014
 Augmented the Deformable Parts Model (DPM) based face detector and landmark estimator with loopy part models. Utilized dual decomposition and an augmented lagrangian technique called ADMM (Alternating Direction Method of Multipliers) to solve the resulting inference problem efficiently, often achieving zero primal dual gap. Applied the model to get results comparable to the state of the art for detection and landmark localization
- **Understanding and Predicting Importance** *Virginia Tech*
 Advisor: Devi Parikh Spring 2014
 Formulated importance prediction in abstract images as a structured prediction problem, where importance is defined as the likelihood of an object in an image being mentioned in a description. Incorporated task related insights into feature and model (structure) selection. Predicted importance of objects at upto 86 % accuracy on the Abstract-50S dataset
- **CanSat 2011** *IIIT Hyderabad*
 Advisor: K.S. Rajan Summer 2010
 Designed, fabricated and launched into the lower space an autonomous micro-satellite carrying a large raw hen's egg intact - from *launch* to *landing*. Ground station set up to monitor the mini-satellite. Involved in CanSat testing, circuitry and sensor integration teams

COURSEWORK

- **Graduate Coursework:** Computer Vision Systems, Advanced Computer Vision, Introduction to Machine Learning, Probabilistic Graphical Models, Independent Study - Deep Learning, Numerical Analysis and Software, Data Analytics-2, Convex Optimization, Deep Learning for Perception, Bayesian Statistics, Mathematical Foundations of Machine Learning, Computability and Algorithms
- **Selected Undergraduate Coursework:** Mobile Robotics, Artificial Neural Networks, Speech Signal Processing, Medical Image Processing, Engineering Systems, Data Structures, Operating Systems and Algorithms

SKILLS

- **Programming Languages:** Python, C++, Lua, Matlab
- **Libraries:** Torch, Tensorflow, Caffe, NLTK (Natural Language ToolKit)
- **Human Computation:** Amazon Mechanical Turk

EXTRA CURRICULAR

- Volunteered in organizing Mid-Atlantic Computer Vision (MACV) workshop at Virginia Tech
- Regular participation in Computer Vision and Machine Learning Reading Group at Virginia Tech

- Hosted all the Talks at Felicity - 2011, annual college fest of IIIT Hyderabad
- Coordinator and Founder- Entrepreneurship Cell at IIIT Hyderabad
- Class Representative for ECE Undergraduate batch
- Member, Students Parliament (Monsoon 2012 and Spring 2013)
- Campus Ambassador for Teach for India at IIIT (2011 to 2012)
- Trained in Carnatic Classical music for 7 years

REFERENCES

- Prof. Devi Parikh, *Georgia Tech* (email: parikh@gatech.edu)
- Dr. Kevin P. Murphy, *Google Research* (email: kpmurphy@google.com)
- Dr. Gal Chechik, *Google Research* (email: gal@google.com)
- Dr. C. Lawrence Zitnick, *Facebook AI Research* (email: zitnick@fb.com)
- Prof. Dhruv Batra, *Georgia Tech* (email: dbatra@gatech.edu)