Shanmukha Ramakrishna Vedantam

1207, Snyder Lane, Apt 1200B, Blacksburg, VA, USA

+1-334-329-9834 • vrama91@vt.edu • https://vrama91.github.io/

Education

- **Virginia Polytechnic Institute and State University, United States (2013-present)**
  Ph.D, Computer Engineering
  Advisor: Prof. Devi Parikh
  Specialization: Computer Vision and Machine Learning
- **Virginia Polytechnic Institute and State University, United States (2013-2016)**
  M.S., Computer Engineering
  GPA: 3.89/4.0, 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.
  GPA: 8.21/10, Advisor: Prof. K. Madhava Krishna
  Specialization: Vision for Robotics

Publications

1. **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*

2. **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 (Short)*

3. **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)**

4. **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)**

5. **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

6. **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*

7. **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*

8. **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

- **Generative Models of Visually Grounded Imagination**: Ramakrishna Vedantam, Ian Fischer, Jonathan Huang, Kevin Murphy. [May, 2017]

- **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]

WORK EXPERIENCE
___

Graduate Research Assistant, Fall 2014 - Present
Bradley Department of ECE at Virginia Tech, United States
- Language, Vision and Common Sense
- Supervisor: Prof. Devi Parikh, Virginia Tech

Research Internship, Summer 2017
Facebook AI Research (FAIR), Menlo Park, CA
- Supervisor: Dr. 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

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, ICCV 2015, CVPR 2016, ECCV 2016, ACCV 2016, ICVGIP 2016, CVPR 2017, BMVC 2017, ICCV 2017, NIPS 2017
- **Journals:** Reviewer for International Journal of Computer Vision (IJCV), IEEE Transactions on Image Processing, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)

Advising

- Satwik Kottur, Carnegie Mellon University (co-advised with Devi Parikh)
  Summer 2015
- Prithvijit Chattopadhyay, Delhi Technical University (co-advised with Devi Parikh)
  Summer 2015

Honors and Achievements

- Outstanding reviewer at CVPR, 2017
- Awarded travel grant of USD 1000 for CVPR, 2017 under Google's Archimedes program
- Finalist for the Adobe Research Fellowship (2016)
- Was selected for, and attended the International Computer Vision Summer School (ICVSS), (2014)
- Best Discussion Participant Award, Advanced Computer Vision Course, Virginia Tech, (Spring, 2014)
- Mentioned in Dean's List for excellence in academics at IIIT Hyderabad, (Monsoon, 2009 & 2011, Spring 2012)
- Winner of Judges award and Peer award at Siemens CTT Intern Tech Challenge, (2012)
- Our team placed 3rd in global aerospace competition CANSAT organized by NASA, AAS and AIAA, (2011)
- Attained top 20 rank in Regional Mathematics Olympiad Organized by National Board for Higher Mathematics (NBHM) from Gujarat State. Qualified for the Indian National Mathematics Olympiad, (2008)
- Finalist for the Bal Shree honor, conferred by the President of India for outstanding creativity in Science, (2008)
- Awarded Chacha Nehru Scholarship for Artistic and Innovative Excellence from National Council of Educational Research and Training (NCERT), (2008)
- Attained All India Rank 134 in National Science Olympiad, (2006)
- All India Rank 13 in Indian National Cartography Association (INCA) Map Quiz, (2006)

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.

Selected Projects

- **Class Discriminative Explanations**   *Google Research*
  Advisor: Gal Chechik and Samy Bengio   Summer and Fall, 2016
  Formulated the task of explaining why the image contains a particular category as opposed to a distractor category using natural language. Formulated a model that at inference time induces class discrimination by trading off explanation and discrimination. Quantitative studies using ranking revealed the generated descriptions are more class discriminative as opposed to a generative baseline.
- **Learning Common Sense**   *Virginia Tech*
  Advisor: Devi Parikh and C. Lawrence Zitnick   Spring 2015
  Proposed a model to classify common sense tuples such as <man> *looks at* <cake> based on how plausible they are. Showed that better performance can be achieved on tuple classification by combining textual and visual cues. This work is accepted to the IEEE International Conference on computer Vision (ICCV), 2015
- **Evaluation of Image Descriptions**   *Virginia Tech*
  Advisor: Devi Parikh and C. Lawrence Zitnick   Fall 2014
  Worked on evaluation of image description methods. Proposed a novel evaluation protocol, named CIDEr (Consensus-based Image Description Evaluation) that utilizes human consensus. Showed that our protocol attains better agreement with humans than existing choices. This work is accepted to the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015

- **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
- **Adopting Abstract Scenes for Semantic Scene Understanding** *Virginia Tech*
  Advisor: Devi Parikh and C. Lawrence Zitnick *Spring 2014*
  Studied image memorability and object level saliency for abstract scenes made of clipart. Conducted experiments on Amazon Mechanical Turk using a game-like interface to collect image memorability and object-saliency annotations. Correlated these properties to semantic importance of objects, mined using text. This work has been accepted to the IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2015
- **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
- **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, C++
- **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, *Virginia Tech* (email: parikh@vt.edu)

- Dr. C. Lawrence Zitnick, *Facebook AI Research* (email: zitnick@fb.com)
- Prof. Dhruv Batra, *Virginia Tech* (email: dbatra@vt.edu)