



RESEARCH POSITIONS

Research Scientist II August 2017 to Present
School of Interactive Computing
at Georgia Institute of Technology

PAST POSITIONS

Bradley Postdoctoral Associate August 2016 to August 2017
Machine Learning & Perception Group
at Virginia Tech with Dhruv Batra

Research Assistant May 2012 to August 2016
School of Informatics and Computing,
at Indiana University with David Crandall

Visiting Research Assistant August 2015 to November 2015
Machine Learning & Perception Group
at Virginia Tech with Dhruv Batra

Visiting Research Assistant May 2014 to August 2014
INRIA - WILLOW Project
at L'cole Normale Supérieure and UC Berkeley with ef Sivic and Alexei A. Efros.



EDUCATION

Ph.D., Computer Science – Indiana University 2016
Thesis: Data-Driven Computer Vision for Science and the Humanities
Committee: David Crandall (Chair), Chunfeng Huang, Predrag Radivojac, Michael Ryou

M.S., Computer Science – Indiana University 2013

B.S., Computer Science – University of West Florida 2011



HONORS & AWARDS

- **CoRL Best Presentation Finalist** 2018
- **DARPA Riser – Plenary Speaker** 2018
Nominated by DARPA program manager to participate in the DARPA Riser event and was selected from 47 candidates as one of three representative speakers for the plenary session at DARPA's 60th anniversary conference D60.

- **Outstanding Reviewer Awards**
 - IEEE Conference on Computer Vision and Pattern Recognition (CVPR)** 2017
Recognition from areas chairs for quality reviewing (awarded to ~ 8.5% of reviewers).
 - IEEE International Conference on Computer Vision (ICCV)** 2017
Recognition from areas chairs for quality reviewing (awarded to ~ 4.6% of reviewers).
 - Neural Information Processing Systems (NIPS)** 2017-2018
Recognition from areas chairs for quality reviewing (awarded to ~ 3.6% of reviewers).

- **Best Paper Awards**
 - Conference on Empirical Methods in Natural Language Processing** 2017
One of 4 best papers (or top 0.26%) out of 1500 submissions (1466 reviewed, 323 accepted) to EMNLP 2017 which is a top venue for research on AI with natural language capabilities.

CVPR Workshop on Egocentric Vision**2014**

Awarded best papers (or top 7.6%) out of 13 accepted papers to the CVPR Workshop on Egocentric Vision, a core workshop for discussion of egocentric (or first-person) vision.

- **Bradley Postdoctoral Fellowship (Virginia Tech)** **2016**
- **HANDS Travel Award (CVPR)** **2016**
- **Doctoral Consortium Travel Award (ICCV)** **2015**
- **Dissertation Development Award (Indiana University)** **2015**
- **Heidelberg Laureate Forum Selection (HLF Foundation)** **2015**



SELECTED PUBLICATIONS

Peer-Reviewed Conference Papers (acceptance rates typically 2-25%)

1. Sainandan Ramakrishnan, Aishwarya Agrawal, Stefan Lee. Overcoming Language Priors in Visual Question Answering with Adversarial Regularization. *Neural Information Processing Systems (NIPS)*, 2018.
2. Abhishek Das, Georgia Gkioxari, Stefan Lee, Devi Parikh, Dhruv Batra. Neural Modular Control for Embodied Question Answering. *Conference on Robot Learning (CoRL)*, 2018.
3. Jianwei Yang*, Jiasen Lu*, Stefan Lee, Dhruv Batra, Devi Parikh. Learning to Ask Questions to Learn Visual Recognition. *Conference on Robot Learning (CoRL)*, 2018. [[Oral Paper](#)] [[Best Presentation Finalist](#)]
4. Ramprasaath R. Selvaraju, Prithvijit Chattopadhyay, Mohamed Elhoseiny, Tilak Sharma, Dhruv Batra, Devi Parikh, Stefan Lee. Choose Your Neuron: Incorporating Domain Knowledge through Neuron Importance. *European Conference on Computer Vision (ECCV)*, 2018.
5. Jianwei Yang*, Jiasen Lu*, Stefan Lee, Dhruv Batra, Devi Parikh. Graph R-CNN for Scene Graph Generation. *European Conference on Computer Vision (ECCV)*, 2018.
6. Ashwin K Vijayakumar, Stefan Lee, Anitha Kannan, Dhruv Batra. Learn From Your Neighbor: Learning Multi-Modal Distributions from Sparse Annotation. *International Conference on Machine Learning (ICML)*, 2018. [[Oral Paper – Long Talk](#)].
7. Abhishek Das, Samyak Datta, Georgia Gkioxari, Stefan Lee, Devi Parikh, Dhruv Batra. Embodied Question Answering. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. [[Oral Paper](#)].
8. Ashwin Vijayakumar, Michael Cogswell, Ramprasath Selvaraju, Qing Sun, Stefan Lee, David Crandall, and Dhruv Batra. Diverse Beam Search for Improved Description of Complex Scenes. *AAAI Conference on Artificial Intelligence (AAAI)*, 2018.
9. Abhishek Das*, Satwik Kottur*, José M.F. Moura, Stefan Lee, and Dhruv Batra. Learning Cooperative Visual Dialog Agents with Deep Reinforcement Learning. *IEEE International Conference on Computer Vision (ICCV)*, 2017. [[Oral Paper: 45/3220 = 1.4%](#)].
10. Satwik Kottur, José M.F. Moura, Stefan Lee, and Dhruv Batra. Natural Language Does Not Emerge 'Naturally' in Multi-Agent Dialog. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2017. [[Best Short Paper: 4/1500 = 0.26%](#)].
11. Aroma Mahendru*, Viraj Prabhu*, Akrit Mohapatra*, Dhruv Batra, and Stefan Lee. The Promise of Premise: Harnessing Question Premises in Visual Question Answering. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2017.
12. Viraj Prabhu, Prithvijit Chattopadhyay, Deshraj Yadav, Arjun Chandrasekaran, Abhishek Das, Stefan Lee, Dhruv Batra, and Devi Parikh. Evaluating Visual Dialog Agents via Cooperative Human-AI Games. *Proceedings of the Fifth AAAI Conference on Human Computation and Crowdsourcing (HCOMP)*, 2017.
13. Qing Sun, Stefan Lee, and Dhruv Batra. Bidirectional Beam Search: Forward-Backward Inference in Neural Sequence Models for Fill-in-the-Blank Image Captioning. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
14. Stefan Lee, Senthil Purushwalkam, Michael Cogswell, Viresh Ranjan, David J. Crandall, and Dhruv Batra. Stochastic Multiple Choice Learning for Training Diverse Deep Ensembles. *Neural Information Processing Systems (NIPS)*, 2016.
15. Sven Bambach, Stefan Lee, David Crandall, Chen Yu, Lending A Hand: Detecting Hands and Recognizing Activities in Complex Egocentric Interactions. *IEEE International Conference on Computer Vision (ICCV)*, 2015.
16. Stefan Lee, Nicolas Maisonneuve, David Crandall, Josef Sivic, Alexei A. Efros. Linking Past to Present: Discovering

- Style in Two Centuries of Architecture. *IEEE International Conference on Computational Photography (ICCP)*, 2015.
17. Stefan Lee, Haipeng Zhang, David Crandall. Predicting Geo-informative Attributes in Large-scale Image Collections using Convolutional Neural Networks. *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2015.
18. Stefan Lee, Sven Bambach, David Crandall, John Franchak, and Chen Yu. This Hand Is My Hand: A Probabilistic Approach to Hand Disambiguation in Egocentric Video. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Workshop on Egocentric Vision*, 2014. **[Best Paper: 1/13 = 7.6%]**.
19. Stefan Lee, Jerome Mitchell, David Crandall, and Geoffery Fox. Estimating Bedrock and Surface Layer Boundaries And Confidence Intervals In Ice Sheet Radar Imagery Using MCMC. *International Conference on Image Processing (ICIP)*, 2014.

Journals

20. Abhishek Das, Satwik Kottur, Khushi Gupta, Avi Singh, Deshraj Yadav, Stefan Lee, José M. F. Moura, Devi Parikh, and DhruvBatra. Visual Dialog. *Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 2018.

Book Chapters

21. David J. Crandall, Yunpeng Li, Stefan Lee, and Daniel P. Huttenlocher. Recognizing Landmarks in Large-Scale Social Image Collections. Large-Scale Visual Geo-Localization. Ed. Amir R. Zamir, Asaad Hakeem, Luc Van Gool, Mubarak Shah, Richard Szeliski. Springer, 2016.

Extended Abstracts / Technical Reports

22. Amit Raj, Cusuh Ham, Huda Alamri, Vincent Cartillier, Stefan Lee, James Hays. *Compositional Generation of Images. NIPS Workshop on Visually-Grounded Interaction and Language*, 2017.
23. Ashwin K Vijayakumar, Michael Cogswell, Ramprasath R. Selvaraju, Qing Sun, Stefan Lee, David Crandall, Dhruv Batra. Diverse Beam Search: Decoding Diverse Solutions from Neural Sequence Models. arXiv, 2016.
24. Sven Bambach, Stefan Lee, David Crandall, John Franchak, Chen Yu. Tracking Hands of Interacting People in Egocentric Video. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Workshop on Observing and Understanding Hands in Action*, 2015.
25. Stefan Lee, Senthil Purushwalkam, Michael Cogswell, David J. Crandall, Dhruv Batra. Why M Heads are Better than One: Training a Diverse Ensemble of Deep Networks. arXiv, 2015.
26. Stefan Lee and David Crandall. Learning to Identify Local Floral with Human Feedback. *IEEE Conference on Computer Vision and Pattern Recognition, Workshop on Computer Vision and Human Computation*, 2014.



TEACHING

Georgia Institute of Technology

CS8903 - Special Problems

Fall 2017 - Present

Mentoring 10 PhD, masters, and undergraduate students: Jingjing Pan, Madhuri Shanbhogue, Aditya Rajagopal, Sainandan Ramakrishnan, Anmol Kalia, Saurabh Kumar, Amit Raj, Arda Perkins, Vineet Vinayak, & Raj Amanabrolu

Virginia Tech

ECE5424 - Introduction to Machine Learning

Fall 2016

Indiana University

B659 - Image Processing and Recognition (Assistant Instructor)

Spring 2015

I399 - Research Methods for Informatics and Computing (Mentor)

Fall 2013

C211 - Introduction to CS (Assistant Instructor)

Fall 2011 - Spring 2013



INVITED TALKS

Training Embodied Agents in Semantically and Perceptually Rich Simulations

DARPA 60th Anniversary (D60) - DARPA Riser Plenary	2018
Towards Goal-Driven, Visually Grounded Dialog Agents	
Carnegie Mellon University - VASC Seminar	2018
Facebook AI Research (FAIR) - AI Seminar	2018
University of Trento - DISI Seminar	2017
Georgia Tech - Deep Learning Course Guest Speaker	2017
Training Diverse Deep Ensembles	
Diversity Meets Deep Networks - CVPR Tutorial	2016



MEDIA COVERAGE

Facebook helped create an AI scavenger hunt that could lead to the first useful home robots - MIT Technology Review 2018

How A Virtual Scavenger Hunt Could Train Robots To Find Things In Your Home - FastCompany 2018

Facebook is training AI to answer questions like humans do – Digital Journal 2018

Research Scientist, Assistant Professor Represent IC in DARPA Risers Event - ML@GT Blog 2018

What is Graph R-CNN? - ML@GT Blog 2018

Choose Your Neuron: Incorporating Domain Knowledge through Neuron-Importance - ML@GT Blog 2018

Embodied Question Answering - ML@GT Blog 2018



SERVICE

Regularly review or serve on the program committee for

Computer Vision and Pattern Recognition (CVPR)

European Conference on Computer Vision (ECCV)

International Conference on Computer Vision (ICCV)

Neural Information Processing Systems (NIPS)

International Conference on Learning Representations (ICLR)

Social Network Analysis and Mining (Springer Journal)

Image and Vision Computing (Elsevier Journal)

Workshop / Tutorial Organization

Visually-Grounded Interaction and Language (ViGIL) Workshop @ NIPS 2018

Visual Learning and Embodied Agents in Simulation Environments (VLEASE) Workshop @ ECCV 2018

Shortcomings in Vision and Language (SiVL) Workshop @ ECCV 2018

VQA Challenge and Visual Dialog Workshop @ CVPR 2018

Visually-Grounded Interaction and Language (ViGIL) Workshop @ NIPS 2017

Diversity Meets Deep Networks Tutorial @ CVPR 2016