



RESEARCH POSITIONS

Assistant Professor September 2019 to Present
School of Electrical Engineering and Computer Science
at Oregon State University

PAST POSITIONS

Visiting Research Scientist Summer 2020
at Facebook AI Research

Research Scientist II August 2017 to August 2019
School of Interactive Computing
at Georgia Institute of Technology

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'École Normale Supérieure and UC Berkley 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 Ryo

M.S., Computer Science – Indiana University 2013

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



HONORS & AWARDS

- **Outstanding Research Scientist – College of Computing, Georgia Tech** 2019
- **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)** 2017,2019,2020
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 (NeurIPS)	2017-2018
Recognition from areas chairs for quality reviewing (awarded to ~ 3.6% of reviewers).	
International Conference on Learning Representations (ICLR)	2018-2019
European Conference on Computer Vision (ECCV)	2020

– **Best Paper Awards**

Conference on Empirical Methods in Natural Language Processing (EMNLP) **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 20-25%)

1. Jacob Krantz, Aaron Gokaslan, Dhruv Batra, Stefan Lee, Oleksandr Maksymets. Waypoint Models for Instruction-guided Navigation in Continuous Environments. *International Conference on Computer Vision (ICCV), 2021*. [[Oral Paper: 210/6236=3.3% of submissions](#)]
2. Oleksandr Maksymets, Vincent Cartillier, Aaron Gokaslan, Erik Wijmans, Wojciech Galuba, Stefan Lee, Dhruv Batra. THDA: Treasure Hunt Data Augmentation for Semantic Navigation. *International Conference on Computer Vision (ICCV), 2021*.
3. Aayam Shrestha, Stefan Lee, Prasad Tadepalli, Alan Fern. DeepAveragers: Offline Reinforcement Learning by Solving Derived Non-Parametric MDPs. *International Conference on Learning Representations (ICLR), 2021*. [[Spotlight Paper: 114/2997=3.8% of submissions](#)]
4. Vincent Cartillier, Zhile Ren, Neha Jain, Stefan Lee, Irfan Essa, Dhruv Batra. Semantic MapNet: Building Allocentric SemanticMaps and Representations from Egocentric Views. *AAAI Conference on Artificial Intelligence (AAAI), 2021*.
5. Saurabh Satish Desai, Stefan Lee. Auxiliary Tasks for Efficient Learning of Point-Goal Navigation. *Winter Conference on Applications of Computer Vision (WACV), 2021*.
6. Peter Anderson, Ayush Shrivastava, Joanne Truong, Arjun Majumdar, Devi Parikh, Dhruv Batra, Stefan Lee. Sim-to-Real Transfer for Vision-and-Language Navigation. *Conference on Robot Learning (CoRL), 2020*.
7. Samyak Datta, Oleksandr Maksymets, Judy Hoffman, Stefan Lee, Dhruv Batra, Devi Parikh. Integrating Egocentric Localization for More Realistic Point-Goal Navigation Agents. *Conference on Robot Learning (CoRL), 2020*.
8. Simon Stepputtis, Joe Campbell, Mariano Phiellipp, Stefan Lee, Chitta Baral, Heni Ben Amor. Language-Conditioned Imitation Learning for Robot Manipulation Tasks. *Neural Information Processing Systems (NeurIPS), 2020*. [[Spotlight Paper: 280/9467=3.0% of submissions](#)]
9. Michael Cogswell, Jiasen Lu, Rishabh Jain, Stefan Lee, Devi Parikh, Dhruv Batra. Dialog without Dialog Data: Learning Visual Dialog Agents from VQA Data. *Neural Information Processing Systems (NeurIPS), 2020*.
10. Meera Hahn, Jacob Krantz, Dhruv Batra, Devi Parikh, James M. Rehg, Stefan Lee, Peter Anderson. Where Are You? Localization from Embodied Dialog. *Empirical Methods in Natural Language Processing (EMNLP), 2020*.
11. Yilin Yang, Longyue Wang, Shuming Shi, Prasad Tadepalli, Stefan Lee, Zhaopeng Tu. On the Sub-Layer Functionalities of Transformer Decoder. *Findings of Empirical Methods in Natural Language Processing (EMNLP), 2020*.
12. Jacob Krantz, Erik Wijmans, Arjun Majumdar, Dhruv Batra, Stefan Lee. Beyond the Nav-Graph: Vision and Language Navigation in Continuous Environments. *European Conference on Computer Vision (ECCV), 2020*.
13. Arjun Majumdar, Ayush Shrivastava, Stefan Lee, Peter Anderson, Devi Parikh, Dhruv Batra. Improving Vision-and-Language Navigation with Image-Text Pairs from the Web. *European Conference on Computer Vision (ECCV), 2020*. [[Spotlight Paper: 161/5025=3.2% of submissions](#)]

14. Abhishek Kadian, Joanne Truong, Aaron Gokaslan, Alexander Clegg, Erik Wijmans, Stefan Lee, Manolis Savva, Sonia Chernova, Dhruv Batra. Are We Making Real Progress in Simulated Environments? Measuring the Sim2Real Gap in Embodied Visual Navigation. *Intelligent Robots and Systems (IROS) / Robotics and Automation Letters (RA-L)*, 2020.
15. Jiasen Lu*, Vedanuj Goswami*, Marcus Rohrbach, Devi Parikh, Stefan Lee. 12-in-1: Multi-Task Vision and Language Representation Learning. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
16. Erik Wijmans, Abhishek Kadian, Ari Morcos, Stefan Lee, Irfan Essa, Devi Parikh, Manolis Savva, Dhruv Batra. DD-PPO: Learning Near-Perfect PointGoal Navigators from 2.5 Billion Frames. *International Conference on Learning Representation (ICLR)*, 2020.
17. Arijit Ray, Karan Sikka, Ajay Divakaran, Stefan Lee, Giedrius Burachas. Sunny and Dark Outside?! Improving Answer Consistency in VQA through Entailed Question Generation. *Empirical Methods in Natural Language Processing (EMNLP)*, 2019.
18. Jiasen Lu, Dhruv Batra, Devi Parikh, Stefan Lee. ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and-Language Tasks. *Neural Information Processing Systems (NeurIPS)*, 2019.
19. Peter Anderson*, Ayush Shrivastava*, Devi Parikh, Dhruv Batra, Stefan Lee. Chasing Ghosts: Instruction Following as Bayesian State Tracking. *Neural Information Processing Systems (NeurIPS)*, 2019.
20. Harsh Agrawal*, Karan Desai*, Xinlei Chen, Rishabh Jain, Dhruv Batra, Devi Parikh, Stefan Lee, Peter Anderson. nocaps: novel object captioning at scale. *IEEE International Conference on Computer Vision (ICCV)*, 2019.
21. Ramprasaath R. Selvaraju, Stefan Lee, Yilin Shen, Hongxia Jin, Dhruv Batra, and Devi Parikh. Taking a HINT: Leveraging Explanations to Make Vision and Language Models More Grounded. *IEEE International Conference on Computer Vision (ICCV)*, 2019.
22. Ashwin Kalyan, Peter Anderson, Stefan Lee, Dhruv Batra. Trainable Decoding of Sets of Sequences for Neural Sequence Models. *International Conference on Machine Learning (ICML)*, 2019.
23. Yash Goyal, Ziyang Wu, Jan Ernst, Dhruv Batra, Devi Parikh, Stefan Lee. Counterfactual Visual Explanations. *International Conference on Machine Learning (ICML)*, 2019.
24. Ramakrishna Vedantam, Karan Desai, Stefan Lee, Marcus Rohrbach, Dhruv Batra, and Devi Parikh. Probabilistic Neural-symbolic Models for Interpretable Visual Question Answering. *International Conference on Machine Learning (ICML)*, 2019.
25. Erik Wijmans*, Samyak Datta*, Oleksandr Maksymets*, Abhishek Das, Georgia Gkioxari, Stefan Lee, Irfan Essa, Devi Parikh, Dhruv Batra. Embodied Question Answering in Photorealistic Environments with Point Cloud Perception. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019. [\[Oral Paper\]](#)
26. Huda Alamri and Vincent Cartillier and Abhishek Das and Jue Wang and Stefan Lee and Peter Anderson and Irfan Essa and Devi Parikh and Dhruv Batra and Anoop Cherian and Tim K. Marks and Chiori Hori. Audio-Visual Scene-Aware Dialog. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
27. Sainandan Ramakrishnan, Aishwarya Agrawal, Stefan Lee. Overcoming Language Priors in Visual Question Answering with Adversarial Regularization. *Neural Information Processing Systems (NeurIPS)*, 2018.
28. Abhishek Das, Georgia Gkioxari, Stefan Lee, Devi Parikh, Dhruv Batra. Neural Modular Control for Embodied Question Answering. *Conference on Robot Learning (CoRL)*, 2018.
29. Jianwei Yang*, Jiasen Lu*, Stefan Lee, Dhruv Batra, Devi Parikh. Learning to Ask Questions to Learn Visual Recognition. *Conference on Robot Learning (CoRL)*, 2018. [\[Oral Paper\]](#) [\[Best Presentation Finalist\]](#)
30. 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.
31. Jianwei Yang*, Jiasen Lu*, Stefan Lee, Dhruv Batra, Devi Parikh. Graph R-CNN for Scene Graph Generation. *European Conference on Computer Vision (ECCV)*, 2018.
32. 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. [\[Oral Paper – Long Talk\]](#).
33. 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. [\[Oral Paper\]](#).
34. Ashwin Vijayakumar, Michael Cogswell, Ramprasaath 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.

35. 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%**].
36. 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%**].
37. 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.
38. 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.
39. 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.
40. 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 (NeurIPS)*, 2016.
41. 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.
42. 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.
43. 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.
44. 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%**].
45. 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

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

Book Chapters

47. 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

48. Sam Greydanus, Stefan Lee, Alan Fern. Piecewise-constant Neural ODEs. arXiv, 2021.
49. Amit Raj, Cusuh Ham, Huda Alamri, Vincent Cartillier, Stefan Lee, James Hays. *Compositional Generation of Images. NeurIPS Workshop on Visually-Grounded Interaction and Language*, 2017.
50. 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.
51. 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.
52. 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.
53. 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

Oregon State University	
CS535 - Deep Learning	Spring 2022
CS434 - Machine Learning and Data Science	Spring/Fall 2021
CS539 - Natural Language Processing with Deep Learning	Winter 2020
CS539 - Embodied AI	Fall 2019
Georgia Institute of Technology	
CS8903 - Special Problems	Fall 2017 - Spring 2019
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

Telling Embodied Agents What To Do (And The Agents Sometimes Doing It In Reality)	
Invited Speaker – Stanford University NLP Seminar Series	2021
Learning Transferable Visiolinguistic Representations	
Invited Speaker – Oregon State AI Seminar Series	2021
Invited Speaker – Arizona State AI Seminar	2021
Invited Speaker – Facebook AI Research	2020
Exploring the Sim-2-Real Gap in Vision-and-Language Navigation	
Invited Speaker – Lantern Workshop @ COLING	2020
Questions at the Intersection of Emergent Communication and Natural Language	
Invited Speaker – Emergent Communication Workshop @ NeurIPS	2019
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



RECENT MEDIA COVERAGE

- Facebook, Georgia Tech & OSU ViLBERT Achieves SOTA on Vision-and-Language Tasks – Medium 2019
- 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

Served as area chair for

Computer Vision and Pattern Recognition (CVPR), 2021-2022

Neural Information Processing Systems (NeurIPS) 2019-2021

International Conference on Machine Learning (ICML) 2020

AAAI Conference on Artificial Intelligence (AAAI) 2019

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 (NeurIPS)

International Conference on Learning Representations (ICLR)

Social Network Analysis and Mining (Springer Journal)

Image and Vision Computing (Elsevier Journal)

Workshop / Tutorial Organization

Embodied AI Workshop @ CVRP 2020-2021

Embodied Agents Workshop @ CVRP 2019

Visually-Grounded Interaction and Language (ViGIL) Workshop @ NeurIPS 2017 / 2018 / 2019 & NAACL 2021

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

Shortcomings in Vision and Language (SiVL) Workshop @ ECCV 2018 / NAACL 2019

VQA Challenge and Visual Dialog Workshop @ CVPR 2018 - 2019

Diversity Meets Deep Networks Tutorial @ CVPR 2016



STUDENTS SUPERVISED

Current Students

Jacob Krantz, Oregon State University [PhD]

Skand Peri, Oregon State University (co-advised with Fuxin Li) [PhD]

Eric Slyman, Oregon State University [PhD]

Zijaio Yang, Oregon State University [PhD]

Xiangshi Shi, Oregon State University [PhD]

Joe Nguyen, Oregon State University [PhD]

Abhinav Jain, Oregon State University (co-advised with Cindy Grimm) [MS]

Alumni

Yilin Yang, Oregon State University [PhD, 2022]

Saurabh Desai, Oregon State University [MS, 2021]

Shridhar Mishra, Oregon State University [MS, 2021]