



# Stefan Lee

Georgia Tech

Monday, March 18, 2019

1:00 pm

Luddy Hall, Rm. 3166

## Human-Towards Intelligent Agents That Can See, Talk, and Act

**Abstract:** For AI agents to fully step into the role of human collaborators, they must be able to perceive their environment and communicate about this understanding with humans in order to coordinate their actions to achieve mutual goals. The development of such holistic agents presents challenging problems for computer vision, natural language processing, and machine learning. Towards this end, I'll discuss a recent line of work developing agents that communicate in natural language regarding visual scenes including both static images and 3D environments. First, I will focus on work developing agents that engage in visually-grounded, question-answer based dialogs -- a task we call Visual Dialog. I will provide an overview of the Visual Dialog task and highlight some challenges faced by deep agents trained for this problem. Then I will discuss follow-up work in which we address some of these challenges by modeling Visual Dialog as a cooperative game between agents in a reinforcement learning setting -- learning dialog agent policies end-to-end, from pixels to multi-agent, multi-round dialog to game reward. Finally, I'll discuss EmbodiedQA, a recent effort to extend beyond static images and ground similar agents into simulated 3D environments.

**Biography:** Stefan Lee is a Research Scientist in the School of Interactive Computing at Georgia Tech where he studies problems at the intersection of machine learning, computer vision, and natural language processing. His current work addresses how to develop agents that can see, talk, and act -- designing agents that can understand and use visually-grounded language to achieve goals in complex environments. His work frequently appears at major conferences in computer vision, natural language processing, and machine learning. He is the recipient of a Best Paper award (EMNLP 2017) and was recognized as a 2018 DARPA Riser for the potential impact of his research agenda. He has also received multiple outstanding reviewer awards (2017 - CVPR, ICCV, NuerIPS. 2018 - NeurIPS, ICLR) recognizing his service efforts in the community. Prior to his current position, he was a Bradley Postdoctoral Fellow at Virginia Tech after receiving his PhD in 2016 from the School of Informatics and Computing at Indiana University advised by David Crandall.

