

**Abstract:**

Building intelligent agents that possess the ability to perceive the rich visual environment around us, communicate this understanding in natural language to humans and other agents, and execute actions in a physical environment, is a long-term goal of Artificial Intelligence. In this talk, I will present some of my recent work at various points on this spectrum in connecting vision and language to actions; from Visual Dialog (CVPR17, ICCV17, HCOMP17) -- where we develop models capable of holding free-form visually-grounded natural language conversation towards a downstream goal and ways to evaluate them -- to Embodied Question Answering (CVPR18) -- where we augment these models to actively navigate in simulated environments and gather visual information necessary for answering questions.

**Bio:**

Abhishek Das is a Computer Science PhD student at Georgia Institute of Technology, advised by Dhruv Batra, and working closely with Devi Parikh. He is interested in deep learning and its applications in building agents that can see (computer vision), think (reasoning & interpretability), talk (language modeling) and act (reinforcement learning). He is a recipient of an Adobe Research Fellowship and a Snap Research Fellowship. He has held internship positions at Queensland Brain Institute (Winter 2013, Winter 2014), Virginia Tech (Fall 2015) and Facebook AI Research (Summer 2017, Winter 2018, Summer 2018). He graduated from Indian Institute of Technology Roorkee in 2015 with a Bachelor's degree in Electrical Engineering.