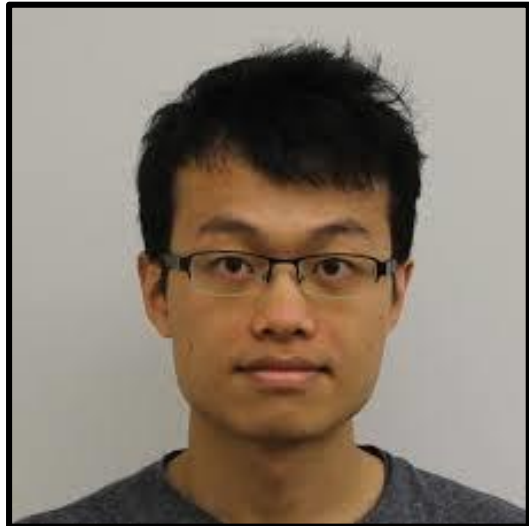




Language
Technologies
Institute

Thesis Proposal

Zoom| Friday, October 30, 2020| 2:00 pm



Joint Analysis and Prediction of Human Actions and Paths in Video

Junwei Liang

Abstract

With the advancement in computer vision deep learning, systems now are able to analyze an unprecedented amount of rich visual information from videos to enable applications such as autonomous driving, socially-aware robot assistant and public safety monitoring. Deciphering human behaviors to predict their future paths/trajectories and what they would do from videos is important in these applications. However, modern vision systems in self-driving applications usually perform the detection (perception) and prediction task in separate components, which leads to error propagation and sub-optimal performance. More importantly, the prediction modules are often limited by the compact interface features extracted by previous perception modules. This design hinders prediction performance in video data from diverse domains and unseen scenarios. To enable optimal future human behavioral forecasting, it is crucial for the system to be able to detect and analyze human activities leading up to the prediction period, passing informative features to the subsequent prediction module. The full vision system should be trained end-to-end to allow gradients to flow through the entire pipeline for optimal performance and generalization ability. In this thesis, with the goal of improving the performance and generalization ability of future trajectory and action prediction models, we conduct human action analysis and jointly optimize models for action detection, prediction and trajectory prediction.

https://drive.google.com/file/d/1OxilbUCti_hyN8dKMEe8oTGeRPFdvDQm/view

COMMITTEE:

**Alexander Haputmann,
(Chair)**

Alan W. Black

Kris Kitani

**Lu Jiang,
(Google Brain)**