MSc Seminar
Thursday October 31, 2019 at 11:00AM in Reynolds, Room 2224

Image Synthesis from Natural Language Descriptions
Arslan Kazmi

Advisor: Dr. Pascal Matsakis
Advisory Committee: Dr. Andrew Hamilton-Wright

ABSTRACT:

Text-to-image synthesis entails the generation of a 2D scene from a natural language text description. This involves taking a text description of an object or a scene, and generating a corresponding 2-dimensional image. A human analogue would be drawing an image entirely from imagination given some caption. Work has been done in the past to facilitate the generation of images from text using natural language processing and rule-based systems. Recently, however, alternate approaches have been posited using machine learning techniques utilizing various deep learning models such as Variational Auto-Encoders and Generative Adversarial Networks. These feature data-driven learning approaches to learn text encodings and their corresponding images, rather than utilizing large object model databases and complex rendering engines.

This seminar aims to discuss the general pipeline (and learning processes) of text-to-image synthesis systems as well as the techniques used at each stage of the pipeline. Each module in the pipeline can be a research area all in its own; it is a challenge to generalize such systems to handle more complex text descriptions.