UNIVERSITY of GUELPH

COLLEGE of ENGINEERING AND PHYSICAL SCIENCES

SCHOOL OF COMPUTER SCIENCE

# MSc Seminar

## Tuesday August 23, 2022 at 12pm online via Zoom


## Giuliano (Julian) Sovernigo


*Preference-based advanced image generation networks and their application in Forensic Art*


**Advisor:** Dr. Xiaodong Lin
**Advisory:** Dr. Andrew Hamilton-Wright


## Abstract:


Forensic art, also known as forensic sketch generation, is an important practice in the field of modern law enforcement. The practice allows investigators to put a face to a crime, warn the public, and focus their efforts on apprehending the correct suspect. Despite its importance, the forensic sketch process remains largely manual, though many other processes have been modernized to include emerging state-of-the-art technology. Currently, the workflow for creating a forensic sketch involves painstaking selection of features from a book of composites by an eyewitness and a forensic artist, requiring hours of back-and-forth work to create a sufficiently accurate composite face. This process is hindered by a large time requirement, a tendency for witnesses to withhold changes they believe to be cumbersome to enact, and a high dependence on the focus and effort of the forensic artist.

Recent breakthroughs in machine learning image generation techniques present a possible solution; the complex and sophisticated architectures proposed by OpenAI's DALL-E2 and Google's Imagen networks in concert with preference-based reinforcement learning (PbRL) could be leveraged to create a framework capable of rapidly generating increasingly specific sketches from user feedback. This new workflow could see the forensic artist able to quickly and effortlessly guide an eyewitness using our tools, freeing both parties from the pitfalls in the current process. To achieve this, our proposed framework will use PbRL techniques on successive sets of user-ranked, generated faces to rapidly narrow the search space for the desired latents for the generated face, which are represented by the learning agent's policy. Since DALL-E2 and Imagen's latter components function as a mapping agent between latent representations (from text) and a space of possible images, the latents optimized through PbRL could then be directly provided to a chosen specialized and adapted generative architecture.

While image generation techniques have been approached using PbRL previously, they have not been focused on the specific generation of composite faces for law enforcement and have been done with relatively simple preference mechanisms and simple generative networks. Our work would be the first to connect the PbRL techniques with the state-of-the-art generative architectures to create a rapid mental-image to realized-face framework. This work will hopefully inspire the creation of new ML-integrated forensic tools capable of reducing the workload required to complete this vital step in an investigation.