Thesis Project Portfolio

An Overview of Facial Recognition Technology

(Technical Report)

Accountability and the Development of Facial Recognition Technology

(STS Research Paper)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Daniel J. McNamara

Spring, 2021

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

(Executive Summary)

An Overview of Facial Recognition Technology and Its Place in Society

The concept of a computer that recognizes human faces sounds like it is straight out of a dystopian science-fiction novel. This is not without good reason, since facial recognition technology (FRT) promises great benefit to society, but inherently carries several ethical questions regarding its use and development. Privacy issues are a primary concern, especially since governments are some of the largest consumers of FRT, but an equally troubling problem with the development and widespread adoption is algorithmic bias. My STS and technical literature reviews combine to present a broad overview of these issues and the potential benefits of the technology.

My technical literature review begins with deep-learning FRT algorithms that leverage the power of machine learning. This requires training with a dataset that teach the algorithms the common patterns that constitute the features of a human face. This powerful method is complex and, implemented correctly, can help create highly accurate and useful FRT systems, such as a system that allows doctors to remotely diagnose eye diseases such as cataracts. However, the process depends heavily on the dataset that is provided by the developers during training. Since this training data is inherent to the algorithm, any biases in the training data will be reflected in the algorithm. Another issue of deep machine learning that is more specific to FRT is the privacy of the subjects of the pictures in the training dataset. These deep-learning models, despite being very complex, can still be vulnerable to model inversion attacks that allow an attacker to generate an approximate copy of images in the training dataset.

The STS section of the literature review looks into the sources and effects of these biases, while considering the intentions and responsibilities of developers and consumers of FRT. A significant portion of the research focuses on the implementation of FRT by the government of India. What began as a tool to find missing children and identify bodies is now being used by Indian law enforcement to identify and track peaceful protestors. This presents an ethical conundrum, since these use cases have much different intentions and stakes. As a hypothetical example, if the algorithm is 90% accurate, then the 10% error rate is easily outweighed by the benefit of potentially identifying a missing child. But when applied to identifying a criminal, the 10% error rate is a shockingly high number, especially when considering potential biases. These algorithmic biases, especially racial biases, become increasingly concerning when applied by law enforcement since the algorithm can serve to compound on existing societal biases.

The STS and technical portions of my research constitute a general overview of FRT and its implications. The technical review identifies some uses and vulnerabilities of FRT systems, while also touching on some of the realities of current and future implementations, including the use of FRT by the United States Customs and Border Protection agency. This is closely linked, even overlapped by, the STS research which focuses on the biases inherent to FRT and how they can potentially compound or expand on existing structural biases in society. The future of the implementation of FRT is murky and riddled with ethical questions, but its continued existence and use seems to be certain. It is important that engineers working on FRT consider existing biases in society, so that they can better understand how their work exists within a larger sociotechnical system.