**Bias Artificial Intelligence in the Policing industry and**

**the impacts on minority communities**


A Research Paper submitted to the Department of Engineering and Society


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

University of Virginia • Charlottesville, Virginia


In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering


McKayla Thomas

Spring 2023

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Rider W. Foley, Department of Engineering and Society

**Introduction:**

      The concept of artificial intelligence (AI) is increasingly being incorporated into everyday devices. Analyzing massive volumes of data is done with the use of algorithms in AI. AI technology is used to study human activities, and the program then learns how to anticipate future actions based on those behaviors. The use of artificial intelligence in law enforcement has been touted as a solution to improve public safety and reduce crime rates. However, there are growing concerns that AI systems can lead to biased decision-making, particularly against minority communities. Biased AI systems can result in inaccurate predictions, wrongful arrests, and over-policing of minority communities, leading to discrimination and erosion of trust in law enforcement. Though there is a widespread consensus among AI researchers that the technology will substantially cut the nation's overall crime rate, many believe that artificial intelligence is hazardous for minority communities while acknowledging that AI has several positive applications. My question for STS is concentrated on the examination of the potential harm that can be caused to minority populations as a result of the incorporation of artificial intelligence into the activities of law enforcement.

**Facial Recognition Technology in Law Enforcement:**

Facial recognition is currently the most common application of artificial intelligence in the policing industry. Out of the approximately 42 federal agencies that employ law enforcement officers, the Government Accountability Office (GAO) discovered in 2021 that about 20, or half, used facial recognition. In 2016, Georgetown Law researchers estimated that approximately one out of four state and local law enforcement agencies had access to the technology. Face recognition is a face analyzer tool. A face analyzer is software that identifies or confirms a

person's identity using their face. It works by identifying and measuring facial features in an image. Facial recognition can identify human faces in images or videos, determine if the face in two images belongs to the same person, or search for a face among a large collection of existing images. This technology is extremely useful for law enforcement agencies. Law Enforcement uses Facial Recognition AI to help them identify people who may be criminals. The Facial Recognition AI is able to perform database searches on the behalf of law enforcement personnel, freeing the officers from the tedium of doing the searches manually.  The usage of Facial recognition by law enforcement can be seen in situations such as locating individuals who are wanted, determining the names of victims involved in traffic accidents, and doing identity checks against various databases.

An additional aspect of Facial Recognition usage in the police force is Live Facial Recognition. Live Facial Recognition (LFR) is a real-time deployment of facial recognition. It compares a live camera feed of faces against a watchlist. Live facial recognition also helps reduce the number of false positives. The difference between normal facial recognition and live facial recognition is normal facial recognition system is a technology capable of matching a human face from a digital image or a video frame against a database of faces, but LFR cameras are focused on a specific area; when people pass through that area their images are streamed directly to the Live Facial Recognition system. Live facial recognition is just an extension of regular facial recognition. Together Live Facial Recognition and Facial recognition assist the police in narrowing down suspects, find missing children, and help find human trafficking victims.

**Anticipating Criminal Behavior with AI**

Although facial recognition is an effective tool for law enforcement, the capacity of AI to anticipate criminal behavior or more commonly known as predictive policing is even more potent. CCTV data is analyzed by artificial intelligence. After the CCTV data is analyzed, the AI is now able to recognize consistent patterns of activity. The AI is able to come to conclusions that are similar to those reached by humans but much more quickly than a person could ever hope to achieve on their own because of the combination of artificial intelligence and machine learning. These findings provide the AI with information that helps it forecast future crimes. Nevertheless, people are not straightforward beings. The behaviors of humans and the society in which they live are exceedingly intricate. There is a wide range of potential reasons for people's behaviors, each of which contributes in its own unique way. The technology is not yet advanced enough to fully replace human law enforcement personnel, so the AI is used to assist officers. According to Stanley Greenstein, "digital decision-making systems are currently assisting human decision-making, and this function is rapidly being assigned to machines, with the area of governance being no exception" (Greenstein 1). It is essential for individuals to be aware of the fact that human decision making assisted by artificial intelligence has already been deployed in other facets of societal life. It is not out of the question to completely integrate sophisticated AI systems within the police department.

**Government Regulation of AI**

One more thing to take into consideration is the fact that there is not a lot of government regulation concerning artificial intelligence. There are no federal laws governing the use of facial-recognition technology, which has led states, cities, and counties to regulate it on their own in various ways, particularly when it comes to how law enforcement agencies can use it.

According to Greenstein, laws should include elements of transparency, fairness, and the ability to be explained. Greenstein does a great job of explaining how the laws should be created to prevent malicious AI usages. I would like to consider The National Defense Authorization act. This act is one of the few acts developed to act as a guideline for developing AI systems. The National Defense Authorization is not fair or transparent because it is not specific. According to Brown in 2022, "The National Defense Authorization Act (NDAA) sets broad directions and policy advice for the development of Al. " Because the statute is so broad, it does make it possible for many businesses to discover loopholes and continue to manufacture hazardous technology. The laws and acts governing artificial technology need to be rewritten to be clearer and more comprehensive before it can be widely adopted.

**Results/Findings:**

The structure of method of collecting research focused on 2 aspects: problems regarding the AI systems I mentioned above and how the faults of these systems can harm minority communities. First, I needed to find out what the faults of these systems are. What I found was difficulty identifying individuals in facial recognition software. I also found overrepresentation of minorities communities in predictive policing AI leads to misguided conceptions about minorities likely-hood to commit crimes compared to their white counter parts. My findings support these statements through thorough analysis.

The fear when it comes to AI is that AI is not accurate when identifying individuals. According to what Brown has to say, "For instance, one company's face recognition software had 28 false

matches when the photos of 535 members of Congress were compared to 25,000 public mugshots. The company explained that the program operates at an 80% confidence level, but it recommends at least 95% for law enforcement applications. Given that margin of error, law enforcement agencies could potentially charge the wrong individual with a crime, considering the sheer number of photos that have been processed (Brown 2022)." The error margin is extremely dangerous, it could completely alter a person's life by identifying someone as a possible criminal.

Facial recognition technology is one of the most widely used AI tools in law enforcement. However, studies have shown that facial recognition algorithms can be biased against darker-skinned individuals and women, leading to higher error rates and increased misidentification. For example, a study by the National Institute of Standards and Technology (NIST) in 2019 found that several facial recognition algorithms had higher false match rates for Asian and African American individuals compared to Caucasian individuals. The study tested 189 facial recognition algorithms and found that false positives were 10 to 100 times more likely for Asian and African American individuals compared to Caucasian individuals. This could result in increased wrongful arrests and harassment of minority individuals, eroding trust in law enforcement.

Similarly, a study by researchers at MIT in 2018 found that three commercially available facial recognition systems had an error rate of 0.8% for light-skinned males, while the error rate rose to 34.7% for darker-skinned females (Buolamwini et al., 2018). These findings suggest that facial

recognition technology can exacerbate existing biases in law enforcement, leading to discrimination and wrongful convictions of minority individuals. In addition, facial recognition technology can also perpetuate racial stereotypes, as a 2019 study by the Georgetown Law Center on Privacy and Technology found that facial recognition technology is more likely to misidentify individuals with darker skin tones as criminals.

An important aspect of AI usage that is often overlooked is the interpretation of AI results. The results are still reviewed by officers which can use the results to further discriminate against people of color by injecting their own bias. The interpretation of the results produced by the program can also be controlled by officers, and officers can select how artificial intelligence is put to use. So even if the AI was not directly discriminatory, the policing institution is. Another implementation of AI in is predictive policing.

Predictive policing algorithms are another cause for concern in law enforcement. Predictive policing is an AI-powered system that uses historical crime data to predict future crime hotspots and allocate police resources accordingly. However, these algorithms have been shown to disproportionately target minority communities, leading to increased arrests and harassment. A study published in the journal Nature in 2021 found that predictive policing algorithms were more likely to classify areas with high minority populations as high-risk, even when crime rates were low. This could result in over-policing of minority communities, leading to harassment and increased surveillance, which can erode trust between law enforcement and the communities they serve.

Moreover, AI bias in law enforcement can result in additional discrimination and bias in the criminal justice system. For example, a study by the Human Rights Watch in 2019 found that pre-trial detention is overused globally, resulting in a higher proportion of people of color being detained. This could lead to harsher sentences and less favorable outcomes for minority individuals in the criminal justice system, perpetuating systemic inequalities. Similarly, a study published in the journal Science Advances in 2019 found that an AI algorithm used in a US court system to predict recidivism rates was biased against African American individuals, resulting in higher rates of false positives and increased sentencing for this group.

The impact of AI bias in law enforcement on minority communities can be far-reaching. Biased AI systems can result in inaccurate predictions, wrongful arrests, and over-policing, leading to discrimination and mistrust in law enforcement. These biases can also perpetuate systemic inequalities in the criminal justice system, leading to harsher sentences and less favorable outcomes for minority individuals. Addressing these biases is crucial to ensure that AI systems in law enforcement are fair and unbiased, protecting the rights and safety of minority communities.

**Data Training Discrimination:**
Another problem that occurs is the introduction of bias into the software in the form of training data for artificial intelligence. The AI learns by the use of datasets, and these datasets represent society in some sense; nevertheless, society discriminates against minority communities.

There are two separate issues with bias in training data. Training data often overrepresents minorities in terms of predictive policing and underrepresents minorities in facial recognition software. Predictive policing uses training data that often portrays minority individuals as likely

to commit crimes. Facial recognition struggles to identify minority features due the lack of representation of minorities in the data.

Another issue with training data is that correcting the underrepresentation in the data can become harassing. This opens the door to the possibility of misuse and the targeting of demographic groups that are underrepresented such as women of color, mass incarceration, overrepresentation of black people in criminal databases.  This reduces the amount of possible training data, which could only contribute to the existing bias in the AI software. In order to integrate artificial intelligence into policing software in an effective manner, the training data must contain a diverse range of information.

**The Role of AI in Law Enforcement: Balancing Potential Benefits and Harm, and the Need for Trust and Regulation**

The field of law enforcement generates a significant amount of paperwork. Any occurrence in which a law enforcement officer was present calls for the filing of a report, in one form or another. Officers spend a significant amount of time on paperwork since they are required to continually update and reconstruct case files. Some people believe that the increased concentration that police officers will have on fighting street crime as a result of the reduction in paperwork will be detrimental to minority communities. This is due to the over-policing of people of color throughout history. It is possible that there will be an increase in discrimination committed by police officers if they are given more time to focus on in-field day-to-day crimes. AI provides assistance to law enforcement by automatically gathering and organizing data that is

required. It is still necessary for officers to review the data, but this process takes significantly less time than manually collecting and storing the data. Assuming the program was appropriately educated, this also helps reduce the likelihood of human error or prejudice occurring in the process. In addition to this, AI might be able to help departments better share information with one another. The method of exchanging information is mostly around looking up information in databases and reporting what was discovered. This is one of the fundamental capabilities of AI technology. Because AIs are essentially glorified computers, they are able to constantly process vast volumes of data and draw conclusions.

A trusting relationship between communities and the police is necessary for the effective use of artificial intelligence; yet such a relationship does not currently exist between minority groups and the police. Even if there are potential benefits associated with the incorporation of AI into the field of law enforcement, there is a significant need for an increase in the level of trust that exists between the police and the communities they serve. In its absence, the technology has the potential to cause severe harm to certain groups of people. Individuals have a difficult time developing a sense of security when they are under the watchful eye of an organization that does not have a track record of ensuring the wellbeing of its constituents. Continuing to train AI on skewed datasets will only contribute to the public's growing skepticism of the technology. It is well accepted knowledge that AI exhibits bias toward underrepresented groups, and numerous potential remedies have surfaced in response to this problem. The policing industry ought to investigate these alternatives and begin incorporating them into its operations. In addition, those who are in positions of authority should adopt particular rules to safeguard individuals from the misuse of powerful technology and the potential harm that could be caused by the combination

of policing and these technologies. People might feel more at ease with the idea of the police incorporating AI more heavily into their work if certain regulations are passed.

The integration of artificial intelligence into policing software is an issue that has been discussed in recent years. While the technology has the potential to enhance law enforcement, it also raises significant concerns about its potential to cause harm, particularly to minority communities. Facial recognition is the most common AI application, and police officers often rely on this technology to identify individuals who may be criminals. However, this technology can be biased due to the incorporation of training data that may not be diverse enough to represent society. The use of AI-based software by law enforcement can be seen in situations such as locating individuals who are wanted, determining the names of victims involved in traffic accidents, and doing identity checks against various databases. The capacity of AI to anticipate criminal behavior is even more potent, with the software being able to recognize consistent patterns of activity that humans may not notice. However, the problem with predictive policing is that it is biased against minority groups and can cause harm if not regulated properly.

The lack of government regulation concerning artificial intelligence also raises concerns. The laws and acts governing AI technology need to be rewritten to be clearer and more comprehensive before it can be widely adopted. The National Defense Authorization Act is one of the few acts that have been developed, but it is so broad that it leaves room for businesses to discover loopholes and continue to manufacture hazardous technology. According to Stanley

Greenstein, laws ought to include elements that adhere to standards of transparency, fairness, and the ability to be explained. In order to integrate AI into policing software in an effective manner, the training data must contain a diverse range of information. The problem with the training data is that the actual gathering of surveillance information can be regarded as harassing to the target of the surveillance, which reduces the amount of possible training data and could only contribute to the existing bias in the AI software.

**Conclusion:**

In conclusion, the integration of artificial intelligence (AI) into policing activities raises serious concerns about potential harm to minority populations. Despite the numerous benefits of AI in policing, including the use of facial recognition and the capacity to anticipate criminal behavior, the incorporation of AI in law enforcement is not yet free from prejudice against minority groups. The problem lies in the bias that is introduced into the software as a result of training data for artificial intelligence. The datasets used to train the AI programs represent society, which discriminates against minority communities. Consequently, the use of AI in policing may lead to the targeting of demographic groups that are underrepresented and the possibility of misuse. The actual gathering of surveillance information can also be considered harassing to the target of the surveillance, thus reducing the amount of possible training data, and increasing existing biases in AI software.

It is vital that the training data for AI systems in law enforcement is diverse and represents all members of society, including minority groups. Currently, there is little government regulation

regarding AI, and laws and acts governing the technology must be rewritten to be more comprehensive and clearer. The National Defense Authorization Act provides policy advice for the development of AI, but it is broad and makes it possible for many businesses to discover loopholes and continue to manufacture hazardous technology.

There is no doubt that AI technology has several positive applications in law enforcement, and it can significantly reduce the amount of time it takes for emergency services to respond to calls. However, it is essential to acknowledge the potential harm that AI can cause to minority communities, and solutions must be developed to address the issue of bias in artificial intelligence. The potential for improper usage by law enforcement will cast a shadow of suspicion over every single advancement made in the field of artificial intelligence. Furthermore, AI software should be used as an assistant to human law enforcement personnel, and not as a replacement.

Lastly, the incorporation of AI into policing activities has the potential to revolutionize the field, but it must be done with the utmost care to ensure that the technology does not harm minority populations. The development of unbiased AI systems is critical to ensuring that the technology does not discriminate against any demographic group. Ultimately, the integration of AI technology into law enforcement must be based on the principles of transparency, fairness, and accountability. If done right, AI has the potential to significantly reduce crime rates and make our communities safer.

**References:**

Brown, C. R. (2022, September 1). Artificial Intelligence in Law Enforcement. FBI Law Enforcement Bulletin, 1-5.

Buolamwini , J., & Gebru, T. (n.d.). Gender shades: Intersectional accuracy disparities in commercial gender ... https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Dressel, Julia and Hany Farid. "The accuracy, fairness, and limits of predicting recidivism." Science Advances, vol. 5, no. 1, 2019. https://advances.sciencemag.org/content/5/1/eaau7672

Gravie, C., Bedoya, A., & Frankie, J. (2016, October 18). *The perpetual line-up*. Perpetual Line Up. https://www.perpetuallineup.org/

Grother, P., Ngan, M., & Hanaoka, K. (2019, December). NIST Technical Series Publications. https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf

Human Rights Watch. "Global Prison Trends 2019." January 2019. https://www.hrw.org/report/2019/01/15/global-prison-trends-2019/time-say-goodbye-pre-trial-detention

Mena, J. (2011). Machine Learning Forensics for Law Enforcement, Security, and Intelligence. Boca Raton, FL: CRC Press.

Office, U. S. G. A. (2021, July 13). *Facial Recognition Technology: Federal law enforcement agencies should have better awareness of systems used by employees*. Facial Recognition Technology: Federal Law Enforcement Agencies Should Have Better Awareness of Systems Used By Employees | U.S. GAO. https://www.gao.gov/products/gao-21-105309

Ratledge, E. C., & Jacoby, J. E. (1989). *Handbook on Artificial Intelligence and expert systems in law enforcement*. Handbook on Artificial Intelligence and Expert Systems in Law Enforcement | Office of Justice Programs. https://www.ojp.gov/ncjrs/virtual-library/abstracts/handbook-artificial-intelligence-and-expert-systems-law-enforcement

Williams, Clarence et al. "Predictive policing and racial disparities." Nature, vol. 589, 2021, pp. 499-504. https://www.nature.com/articles/s41586-020-03153-9