The New York Times | https://www.nytimes.com/2018/06/21/opinion/facial-analysis-technology-bias.html

When the Robot Doesn't See Dark Skin

By Joy Buolamwini
Ms. Buolamwini is the founder of the Algorithmic Justice League.

June 21, 2018

When I was a college student using A.I.-powered facial detection software for a coding project, the robot I programmed couldn't detect my dark-skinned face. I had to borrow my white roommate's face to finish the assignment. Later, working on another project as a graduate student at the M.I.T. Media Lab, I resorted to wearing a white mask to have my presence recognized.

My experience is a reminder that artificial intelligence, often heralded for its potential to change the world, can actually reinforce bias and exclusion, even when it's used in the most well-intended ways.

A.I. systems are shaped by the priorities and prejudices — conscious and unconscious — of the people who design them, a phenomenon that I refer to as "the coded gaze." Research has shown that automated systems that are used to inform decisions about sentencing produce results that are biased against black people and that those used for selecting the targets of online advertising can discriminate based on race and gender.

Specifically, when it comes to algorithmic bias in facial analysis technology — my area of research and one focus of my work with the Algorithmic Justice League — Google's photo application labeling black people in images as "gorillas" and facial analysis software that works well for white men but less so for everyone else are infamous examples. As disturbing as they are, they do not fully capture the risks of this technology that is increasingly being used in law enforcement, border control, school surveillance and hiring.

The products of a company called HireVue, which are used by over 600 companies including Nike, Unilever and even Atlanta Public Schools, allow employers to interview job applicants on camera, using A.I. to rate videos of each candidate according to verbal and nonverbal cues. The company's aim is to reduce bias in hiring.

But there's a catch: The system's ratings, according to a Business Insider reporter who tested the software and discussed the results with HireVue's chief technology officer, reflect the previous preferences of hiring managers. So if more white males with

generally homogeneous mannerisms have been hired in the past, it's possible that algorithms will be trained to favorably rate predominantly fair-skinned, male candidates while penalizing women and people of color who do not exhibit the same verbal and nonverbal cues.

It's repeatedly been proven that apart from technology, people tend to make hiring decisions favoring white and male candidates, all other things being equal. With this in mind, the instinct to hand the rating of potential employees over to technology is understandable. But how do we know a qualified candidate whose verbal and nonverbal cues tied to age, gender, sexual orientation or race depart from those of the high performers used to train the algorithm will not be scored lower than a similar candidate who more closely resemblesthe in-group? We won't know if we do not repeatedly test the technology and its application.

The tests that have been done on facial analysis technology raise concerns. In collaboration with the computer vision expert Timnit Gebru, I investigated the accuracy of facial analysis technology from IBM, Microsoft and Face++. On the simple task of guessing the gender of a face, all companies' technology performed better on male faces than on female faces and especially struggled on the faces of dark-skinned African women. In the worst case, the technology was 34 percent less accurate for those women than it was for white men.

Given how susceptible facial analysis technology seems to recreating gender and racial bias, companies using HireVue, if they hope to increase fairness, should check their systems to make sure it is not amplifying the biases that informed previous hiring decisions. It's possible companies using HireVue could someday face lawsuits charging that the program had a negative disparate impact on women and minority applicants, a violation of Title VII of the Civil Rights Act.

The risks of biased facial analysis technology extend beyond hiring. According to the Center on Privacy and Technology at Georgetown Law, the faces of half of all adults in the United States — over 117 million people — are currently in face recognition database networks that can be searched by police departments without warrant. These searches are often reliant on facial recognition technology that hasn't been tested for accuracy on different groups of people. This matters because misidentification can subject innocent people to police scrutiny or erroneous criminal charges.

In the case of South Wales, where Big Brother Watch reports that between May 2017 and March 2018 the faces of over 2,400 misidentified innocent people were stored by the police department without their consent, the department reported a false-positive facial identification rate of 91 percent. But it's important to remember that even if false-positive match rates improve, unfair use of facial recognition technology cannot be fixed with a software patch. Even accurate facial recognition can be used in disturbing ways. The

Baltimore police department used face recognition technology to identify and arrest people who attended the 2015 protests against police misconduct that followed Freddie Gray's death in Baltimore.

We need to challenge the growing use of this technology, and there has been some progress on this front. The A.C.L.U. is calling on Amazon to stop selling facial analysis technology to law enforcement and is contesting the use of in-car facial recognition for the Vehicle Face System being tested at the United States-Mexico border. Though lawmakers in Texas, Illinois and California have made legislative efforts to regulate facial recognition technology, there are no federal laws. Yet, there is a blueprint. A 2016 report from Georgetown Law School proposed model federal legislation. Policymakers should embrace it.

We can also learn from international models. Unlike the United States, Canada has a federal statute governing the use of biometric data in the private sector. Companies like Facebook and Amazon must obtain informed consent to collect citizens' unique face information. In the European Union, Article 9 of the General Data Protection Regulation requires express affirmative consent for collection of biometrics from E.U. citizens.

Everyday people should support lawmakers, activists and public-interest technologists in demanding transparency, equity and accountability in the use of artificial intelligence that governs our lives. Facial recognition is increasingly penetrating our lives, but there is still time to prevent it from worsening social inequalities. To do that, we must face the coded gaze.