TECHNOLOGY

# Facial-Recognition Software Might Have a Racial Bias Problem

Depending on how algorithms are trained, they could be significantly more accurate when identifying white faces than African American ones.

CLARE GARVIE AND JONATHAN FRANKLE APR 7, 2016

A supervisor with the ID Fraud Unit of the North Carolina Department of Motor Vehicles looks through photos in the facial-recognition system. (GERRY BROOME / AP)

In 16 "undisclosed locations" across northern Los Angeles, digital eyes watch the public. These aren't ordinary police-surveillance cameras;

away. The faces they collect are then compared, in real-time, against "hot lists" of people suspected of gang activity or having an open arrest warrant.

Considering arrest and incarceration rates across L.A., chances are high that those hot lists disproportionately implicate African Americans. And recent research suggests that the algorithms behind facial-recognition technology may perform *worse* on precisely this demographic. Facial-recognition systems are more likely either to misidentify or fail to identify African Americans than other races, errors that could result in innocent citizens being marked as suspects in crimes. And though this technology is being rolled out by law enforcement across the country, little is being done to explore—or correct—for the bias.

State and local police began using facial recognition in the early 2000s. The early systems were notoriously unreliable, but today law-enforcement agencies in Chicago, Dallas, West Virginia, and elsewhere have acquired or are actively considering more sophisticated surveillance camera systems. Some of these systems can capture the faces of passersby and identify them in real-time. Sheriff's departments across Florida and Southern California have been outfitted with smartphone or tablet facial recognition systems that can be used to run drivers and pedestrians against mug shot databases. In fact, Florida and several other states enroll every driver's license photo in their facial recognition databases. Now, with the click of a button, many police departments can identify a suspect caught committing a crime on camera, verify the identity of a driver who does not produce a license, or search a state driver's license database for suspected fugitives.

perfect. Companies market facial recognition technology as "a highly efficient and accurate tool" with "an identification rate above 95 percent." In reality, these claims are almost impossible to verify. The facial-recognition algorithms used by police are not required to undergo public or independent testing to determine accuracy or check for bias before being deployed on everyday citizens. More worrying still, the limited testing that has been done on these systems has uncovered a pattern of racial bias.

The National Institute of Standards and Technologies (NIST) conducts voluntary tests of facial-recognition vendors every four years. In 2010, NIST observed that accuracy rates had improved tenfold between each round of testing, a dramatic testament to the technology's rapid advances.

But research suggests that the improving accuracy rates are not distributed equally. To the contrary, many algorithms display troubling differences in accuracy across race, gender, and other demographics. A 2011 study, co-authored by one of the organizers of NIST's vendor tests, found that algorithms developed in China, Japan, and South Korea recognized East Asian faces far more readily than Caucasians. The reverse was true for algorithms developed in France, Germany, and the United States, which were significantly better at recognizing Caucasian facial characteristics. This suggests that the conditions in which an algorithm is created—particularly the racial makeup of its development team and test photo databases—can influence the accuracy of its results.

Similarly, a study conducted in 2012 that used a collection of mug shots from Pinellas County, Florida to test the algorithms of three police in California, Maryland, Pennsylvania, and elsewhere. The study, co-authored by a senior FBI technologist, found that all three algorithms consistently performed 5-to-10 percent worse on African Americans than on Caucasians. One algorithm, which failed to identify the right person in 1 out of 10 encounters with Caucasian subjects, failed nearly twice as often when the photo was of an African American.

This bias is particularly unsettling in the context of the vast racial disparities that already exist in police traffic stop, stop and frisk, and arrest rates across the country. African Americans are at least twice as likely to be arrested as members of any other race in the United States and, by some estimates, up to 2.5 times more likely to be targeted by police surveillance. This overrepresentation in both mug shot databases and surveillance photos will compound the impact of that 5-to-10 percent difference in accuracy rates. In other words, not only are African Americans more likely to be misidentified by a facial-recognition system, they're also more likely to be enrolled in those systems and be subject to their processing.

Imagine police are investigating a robbery that was caught on camera. When they run a video still of the suspect's face against their facial-recognition database, they receive 10 possible matches, but none are a perfect match to the suspect. Nonetheless, the closest match is treated as a lead, and police begin investigating an innocent person. Thanks to the accuracy-rate bias in facial-recognition algorithms today, this scenario is statistically more likely to happen to an African American than a white person.

This is not to say that facial-recognition algorithms are "racist," or that unintentionally at a number of points in the process of designing and deploying a facial recognition system. The engineer that develops an algorithm may program it to focus on facial features that are more easily distinguishable in some races than in others—the shape of a person's eyes, the width of the nose, the size of the mouth or chin. This decision, in turn, might be based on preexisting biological research about face identification and past practices which themselves may contain bias. Or the engineer may rely on his or her own experience in distinguishing between faces—a process that is influenced by the engineer's own race.

In addition, algorithms learn how to calculate the similarity of photos by practicing on pre-existing training sets of faces. So even if the features on which an algorithm focuses are race-neutral, a training set of images that contains disproportionate numbers of one race will bias the algorithm's accuracy rates in that direction. The 2012 study on mug shots found that an algorithm trained exclusively on either African American or Caucasian faces recognized members of the race in its training set more readily than members of any other race.

Unfortunately, the two studies discussed here are among the only works on racial bias published in the past decade—far too little review for a technology with the power to implicate people as suspects in a criminal investigation. NIST is well placed to lead the development of a comprehensive set of bias tests that could, alongside its existing regime, form the basis of a formal certification framework for facial recognition systems. But NIST testing of facial recognition systems is voluntary. Law-enforcement agencies—or the city councils and state legislatures that pay their bills—are well-placed to require that facial recognition software vendors submit to NIST's existing accuracy tests, Until those requirements are put in place, facial-recognition vendors should voluntarily submit their algorithms to NIST's existing testing regime and other public, peer-reviewed research to measure—and begin to correct—the racial bias in their algorithms.

In the meantime, we're conducting an in-depth study on facial-recognition use by state and local law-enforcement agencies across the country. Our research, based on responses to Freedom of Information requests sent to more than 100 departments, aims to develop a clear picture of what these systems look like, how and on whom they are used, and what policies and legal standards—if any—are in place to constrain their use. The final report, to be released this summer, will include recommendations for government at the federal, state, and local level, the law enforcement agencies themselves, companies, and advocates on how to ensure this technology used in a manner consistent with privacy and civil liberty interests of all citizens, regardless of race.

*We want to hear what you think about this article. Submit a letter to the editor or write to letters@theatlantic.com.*