Emotion-reading tech fails the racial bias test

January 3 2019, by Lauren Rhue

Basketball players Darren Collision (left) and Gordon Hayward (right). Credit: basketball-reference.com

Facial recognition technology has progressed to point where it now interprets emotions in facial expressions. This type of analysis is

increasingly used in daily life. For example, companies can use [facial recognition software](link) to help with hiring decisions. [Other programs](link) scan the faces in crowds to identify threats to public safety.

Unfortunately, this technology struggles to interpret the emotions of black [faces](link). [My new study](link), published last month, shows that [emotional](link) analysis technology assigns more negative emotions to black men's faces than white men's faces.

This isn't the first time that facial [recognition](link) programs have been shown to be biased. Google labeled [black faces as gorillas](link). Cameras identified [Asian faces as blinking](link). Facial recognition programs struggled to correctly [identify gender for people with darker skin](link).

My work contributes to a growing call [to better understand the hidden bias in artificial intelligence software](link).
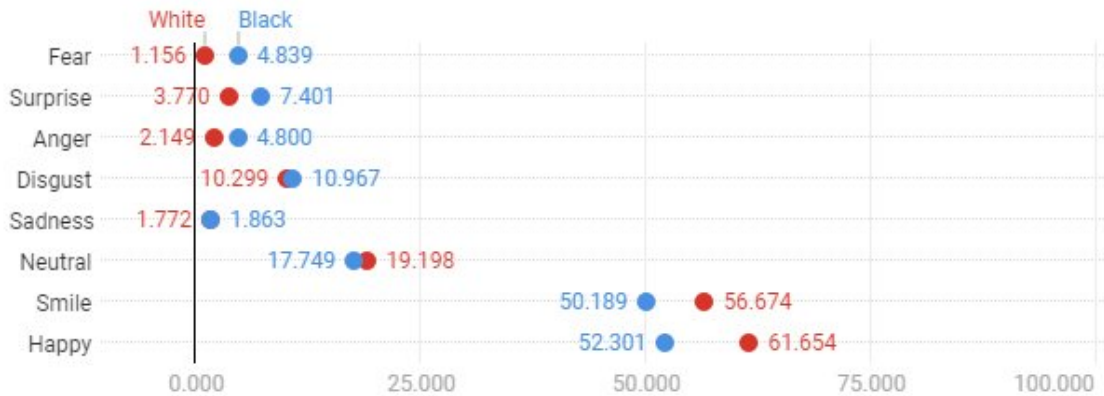
Measuring bias

To examine the bias in the facial recognition systems that analyze people's emotions, I used a data set of 400 NBA player photos from the 2016 to 2017 season, because players are similar in their clothing, athleticism, age and gender. Also, since these are professional portraits, the players look at the camera in the picture.

I ran the images through two well-known types of emotional recognition software. Both assigned black players more negative emotional scores on average, no matter how much they smiled.

Face++

Face++ rated the emotions on facial expressions of basketball players out of 100. Black faces were, on average, rated as angrier and unhappier than white faces.

	White	Black
Fear	1.156	4.839
Surprise	3.770	7.401
Anger	2.149	4.800
Disgust	10.299	10.967
Sadness	1.772	1.863
Neutral	17.749	19.198
Smile	50.189	56.674
Happy	52.301	61.654

For example, consider the official NBA pictures of Darren Collison and Gordon Hayward. Both players are smiling, and, according to the facial recognition and analysis program Face++, Darren Collison and Gordon Hayward have similar smile scores – 48.7 and 48.1 out of 100, respectively.

However, Face++ rates Hayward's expression as 59.7 percent happy and 0.13 percent angry and Collison's expression as 39.2 percent happy and 27 percent angry. Collison is viewed as nearly as angry as he is happy and far angrier than Hayward – despite the facial recognition program itself recognizing that both players are smiling.

In contrast, Microsoft's Face API viewed both men as happy. Still, Collison is viewed as less happy than Hayward, with 98 and 93 percent happiness scores, respectively. Despite his smile, Collison is even scored

with a small amount of contempt, whereas Hayward has none.

Across all the NBA pictures, the same pattern emerges. On average, Face++ rates black faces as twice as angry as white faces. Face API scores black faces as three times more contemptuous than white faces. After matching players based on their smiles, both facial analysis programs are still more likely to assign the negative emotions of anger or contempt to black faces.

Stereotyped by AI

My study shows that facial recognition programs exhibit two distinct types of bias.

Face API

Face API rated the emotions on facial expressions of basketball players out of 100. White faces were seen, on average, as happier than black faces.

Credit: Chart: The Conversation, CC-BY-ND Source: SSRN (2018)

First, black faces were consistently scored as angrier than white faces for every smile. Face++ showed this type of bias. Second, black faces were always scored as angrier if there was any ambiguity about their facial expression. Face API displayed this type of disparity. Even if black faces are partially smiling, my analysis showed that the systems assumed more negative emotions as compared to their white counterparts with similar expressions. The average emotional scores were much closer across races, but there were still noticeable differences for black and white faces.

This observation aligns with other research, which suggests that black professionals must amplify positive emotions to receive parity in their workplace performance evaluations. Studies show that people perceive black men as more physically threatening than white men, even when they are the same size.

Some researchers argue that facial recognition technology is more objective than humans. But my study suggests that facial recognition reflects the same biases that people have. Black men's facial expressions are scored with emotions associated with threatening behaviors more often than white men, even when they are smiling. There is good reason to believe that the use of facial recognition could formalize preexisting stereotypes into algorithms, automatically embedding them into everyday life.

Until facial recognition assesses black and white faces similarly, black people may need to exaggerate their positive facial expressions – essentially smile more – to reduce ambiguity and potentially negative interpretations by the technology.

Although innovative, artificial intelligence can perpetrate and exacerbate existing power dynamics, leading to disparate impact across racial/ethnic groups. Some societal accountability is necessary to ensure fairness to all

groups because facial recognition, like most artificial intelligence, is often invisible to the people most affected by its decisions.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation