

KHARI JOHNSON BUSINESS 05.19.2021 04:23 PM

Twitter's Photo Crop Algorithm Favors White Faces and Women

A study of 10,000 images found bias in what the system chooses to highlight. Twitter has stopped using it on mobile, and will consider ditching it on the web.



PHOTOGRAPH: LORENZO ANTONUCCI/GETTY IMAGES

APPLICATION: CONTENT MODERATION, ETHICS, RECOMMENDATION ALGORITHM

END USER: BIG COMPANY, CONSUMER SECTOR: SOCIAL MEDIA SOURCE DATA: IMAGES

TECHNOLOGY: MACHINE VISION

LAST FALL, CANADIAN student Colin Madland noticed that Twitter's automatic cropping algorithm continually selected his face—not his darker-skinned colleague's—from photos of the pair to display in tweets. The episode ignited accusations of bias as a flurry of Twitter users published elongated photos to see whether the AI would choose the face of a white person over a Black person or if it focused on women's chests over their faces.

At the time, a Twitter spokesperson said assessments of the algorithm before it went live in 2018 found no evidence of race or gender bias. Now, the largest analysis of the AI to date has found the opposite: that Twitter's algorithm favors white people over Black people. That assessment also found that the AI for predicting the most interesting part of a photo does not focus on women's bodies over women's faces.

Previous tests by Twitter and researcher Vinay Prabhu involved a few hundred images or fewer. The analysis released by Twitter research scientists Wednesday is based on 10,000 image pairs of people from different demographic groups to test whom the algorithm favors.

Researchers found bias when the algorithm was shown photos of people from two demographic groups. Ultimately, the algorithm picks one person whose face will appear in Twitter timelines, and some groups are better represented on the platform than others. When researchers fed a picture of a Black man and a white woman into the system, the algorithm chose to display the white woman 64 percent of the time and the Black man only 36 percent of the time, the largest gap for any demographic groups included in the analysis. For images of a white woman and a white man, the algorithm displayed the woman 62 percent of the time. For images of a white woman and a Black woman, the algorithm displayed the white woman 57 percent of the time.

On May 5, Twitter did away with image cropping for single photos posted using the Twitter smartphone app, an approach Twitter chief design officer Dantley Davis favored since the algorithm controversy erupted last fall. The change led people to post tall photos and signaled the end of "open for a surprise" tweets.

The so-called saliency algorithm is still in use on Twitter.com as well as for cropping multi-image tweets and creating image thumbnails. A Twitter spokesperson says excessively tall or

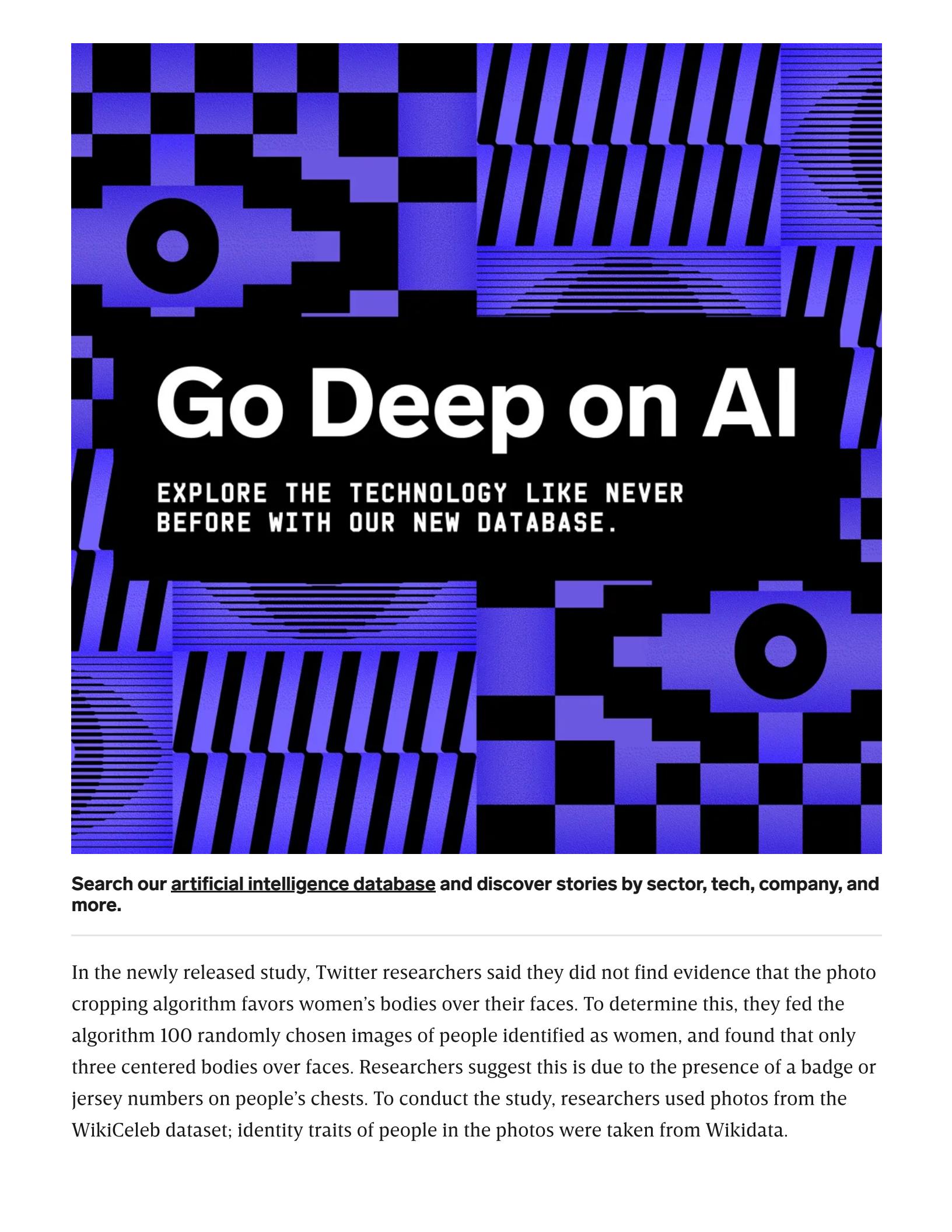
wide photos are now center cropped, and the company plans to end use of the algorithm on the Twitter website. Saliency algorithms are trained by tracking what people look at when they look at an image.

Other sites, including Facebook and Instagram, have used AI-based automated cropping. Facebook did not respond to a request for comment.

Accusations of gender and race bias in computer vision systems are, unfortunately, fairly common. Google recently detailed efforts to improve how Android cameras work for people with dark skin. Last week the group Algorithm Watch found that image-labeling AI used on an iPhone labeled cartoon depictions of people with dark skin as “animal.” An Apple spokesperson declined to comment.

Regardless of the results of fairness measurements, Twitter researchers say algorithmic decisionmaking can take choice away from users and have far-reaching impact, particularly for marginalized groups of people.

Keep Reading



Go Deep on AI

EXPLORE THE TECHNOLOGY LIKE NEVER BEFORE WITH OUR NEW DATABASE.

Search our [artificial intelligence database](#) and discover stories by sector, tech, company, and more.

In the newly released study, Twitter researchers said they did not find evidence that the photo cropping algorithm favors women's bodies over their faces. To determine this, they fed the algorithm 100 randomly chosen images of people identified as women, and found that only three centered bodies over faces. Researchers suggest this is due to the presence of a badge or jersey numbers on people's chests. To conduct the study, researchers used photos from the WikiCeleb dataset; identity traits of people in the photos were taken from Wikidata.

The Twitter paper acknowledges that by limiting the analysis to Black or white or male and female comparisons, it can exclude people who identify as nonbinary or mixed race.

Researchers said they had hoped to use the Gender Shades dataset created to assess the performance of facial recognition systems based on skin tone, but licensing issues arose.

Twitter published the study on the preprint repository arXiv. A Twitter spokesperson said it had been submitted to a research conference to be held in October.

Twitter research scientists suggest that the racial bias found in the analysis may be a result of the fact that many images in the WikiCeleb database have dark backgrounds and the saliency algorithm is drawn to the higher contrast of photos showing people with light skin against a dark background. They also suggest that dark eye color on light skin played a role in saliency algorithms favoring people with light skin.

Coauthors of the paper come from Twitter's ML Ethics, Transparency, and Accountability (META) Team, which Twitter launched last month. Rumman Chowdhury, founder of algorithm auditing startup Parity and a former adviser to tech companies and governments, directs the team.

In a blog post last month, Twitter said it created the team to take responsibility for Twitter's use of algorithms, provide transparency into internal decisionmaking about AI that impacts hundreds of millions of people, and hold the company accountable. Some questions remain about how the META team will operate, such as who makes the final call about whether Twitter uses certain kinds of AI.

The Twitter spokesperson said cross-functional teams decide what actions are taken on algorithms, but did not address the question of who has final authority to decide when a form of AI is considered too unfair for use.

See What's Next in Tech With the Fast Forward Newsletter

From artificial intelligence and self-driving cars to transformed cities and new startups, sign up for the latest news.

Your email	<input type="text"/>
Enter your email	
<input type="button" value="SUBMIT"/>	

By signing up you agree to our [User Agreement](#) and [Privacy Policy & Cookie Statement](#)

In the coming months, META plans to assess how the Twitter home page recommendation algorithms treat certain racial groups and how Twitter's AI treats people based on political ideologies.

The creation of the META team came amid questions about the independence and viability of ethical AI teams in corporate environments. In a move that has since led [AI groups to turn down funding](#) and thousands of Googlers to revolt, Google and former Ethical AI team lead Timnit Gebru parted ways in December 2020. In an [interview](#) shortly after, Chowdhury said the episode has consequences for responsible AI and the entire AI industry.

As Chowdhury [pointed out earlier this year](#), there are many ways to define an algorithm audit. What's not included in Twitter's saliency research: analysis of data used to train Twitter's saliency algorithm, or more detailed information about the analysis Twitter carried out before the saliency algorithm came into use.

When asked how the Twitter saliency controversy changed company policy, a company spokesperson said that the company conducts risk assessments around privacy and security, and that the META team is creating fairness metrics for the company's model experimentation platform and standards for the ethical review of algorithms.

More Great WIRED Stories

-  The latest on tech, science, and more: [Get our newsletters!](#)
- The 60-year-old scientific screwup [that helped Covid kill](#)
- The cicadas are coming. [Let's eat them!](#)
- Decades-old flaws affect [almost every Wi-Fi device](#)
- How to take a slick, professional [headshot with your phone](#)
- What a crossword AI reveals about [humans' way with words](#)
-  Explore AI like never before with [our new database](#)
-  WIRED Games: Get the latest [tips, reviews, and more](#)
-  Optimize your home life with our Gear team's best picks, from [robot vacuums](#) to [affordable mattresses](#) to [smart speakers](#)

ways AI shapes human lives. He was previously a senior writer at VentureBeat, where he wrote stories about power, policy, and novel or noteworthy uses of AI by businesses and governments. He is based...
[Read more](#)

SENIOR WRITER

TOPICS ARTIFICIAL INTELLIGENCE ALGORITHMS MACHINE LEARNING BIAS TWITTER COMPUTER VISION

1 Year of WIRED for
\$10 \$5.
END OF SUMMER SALE

SUBSCRIBE