

The New York Times

Amazon Is Pushing Facial Technology That a Study Says Could Be Biased

In new tests, Amazon's system had more difficulty identifying the gender of female and darker-skinned faces than similar services from IBM and Microsoft.

By **Natasha Singer**

Jan. 24, 2019

Over the last two years, Amazon has aggressively marketed its facial recognition technology to police departments and federal agencies as a service to help law enforcement identify suspects more quickly. It has done so as another tech giant, Microsoft, has called on Congress to regulate the technology, arguing that it is too risky for companies to oversee on their own.

Now a new study from researchers at the M.I.T. Media Lab has found that Amazon's system, Rekognition, had much more difficulty in telling the gender of female faces and of darker-skinned faces in photos than similar services from IBM and Microsoft. The results raise questions about potential bias that could hamper Amazon's drive to popularize the technology.

In the study, published Thursday, Rekognition made no errors in recognizing the gender of lighter-skinned men. But it misclassified women as men 19 percent of the time, the researchers said, and mistook darker-skinned women for men 31 percent of the time. Microsoft's technology mistook darker-skinned women for men just 1.5 percent of the time.

A study published a year ago found similar problems in the programs built by IBM, Microsoft and Megvii, an artificial intelligence company in China known as Face++. Those results set off an outcry that was amplified when a co-author of the study, Joy Buolamwini, posted YouTube videos showing the technology misclassifying famous African-American women, like Michelle Obama, as men.

The companies in last year's report all reacted by quickly releasing more accurate technology. For the latest study, Ms. Buolamwini said, she sent a letter with some preliminary results to Amazon seven months ago. But she said that she hadn't heard back from Amazon, and that when she and a co-author retested the company's product a couple of months later, it had not improved.

Matt Wood, general manager of artificial intelligence at Amazon Web Services, said the researchers had examined facial analysis — a technology that can spot features such as mustaches or expressions such as smiles — and not facial recognition, a technology that can match faces in photos or video stills to identify individuals. Amazon markets both services.

“It’s not possible to draw a conclusion on the accuracy of facial recognition for any use case — including law enforcement — based on results obtained using facial analysis,” Dr. Wood said in a statement. He added that the researchers had not tested the latest version of Rekognition, which was updated in November.

Amazon said that in recent internal tests using an updated version of its service, the company found no difference in accuracy in classifying gender across all ethnicities.



The M.I.T. researchers used these and other photos to study the accuracy of facial technology in identifying gender.

With advancements in artificial intelligence, facial technologies — services that can be used to identify people in crowds, analyze their emotions, or detect their age and facial characteristics — are proliferating. Now, as companies begin to market these services more aggressively for uses like policing and vetting job candidates, they have emerged as a lightning rod in the debate about whether and how Congress should regulate powerful emerging technologies.

The new study, scheduled to be presented Monday at an artificial intelligence and ethics conference in Honolulu, is sure to inflame that argument.

Proponents see facial recognition as an important advance in helping law enforcement agencies catch criminals and find missing children. Some police departments, and the Federal Bureau of Investigation, have tested Amazon's product.

But civil liberties experts warn that it can also be used to secretly identify people — potentially chilling Americans' ability to speak freely or simply go about their business anonymously in public.

Over the last year, Amazon has come under intense scrutiny by federal lawmakers, the American Civil Liberties Union, shareholders, employees and academic researchers for marketing Rekognition to law enforcement agencies. That is partly because, unlike Microsoft, IBM and other tech giants, Amazon has been less willing to publicly discuss concerns.

Amazon, citing customer confidentiality, has also declined to answer questions from federal lawmakers about which government agencies are using Rekognition or how they are using it. The company's responses have further troubled some federal lawmakers.

“Not only do I want to see them address our concerns with the sense of urgency it deserves,” said Representative Jimmy Gomez, a California Democrat who has been investigating Amazon's facial recognition practices. “But I also want to know if law enforcement is using it in ways that violate civil liberties, and what — if any — protections Amazon has built into the technology to protect the rights of our constituents.”

In a letter last month to Mr. Gomez, Amazon said Rekognition customers must abide by Amazon's policies, which require them to comply with civil rights and other laws. But the company said that for privacy reasons it did not audit customers, giving it little insight into how its product is being used.

The study published last year reported that Microsoft had a perfect score in identifying the gender of lighter-skinned men in a photo database, but that it misclassified darker-skinned women as men about one in five times. IBM and Face++ had an even higher error rate, each misclassifying the gender of darker-skinned women about one in three times.

Ms. Buolamwini said she had developed her methodology with the idea of harnessing public pressure, and market competition, to push companies to fix biases in their software that could pose serious risks to people.

Ms. Buolamwini, who had done similar tests last year, conducted another round to learn whether industry practices had changed, she said. Tony Luong for The New York Times

“One of the things we were trying to explore with the paper was how to galvanize action,” Ms. Buolamwini said.

Immediately after the study came out last year, IBM published a blog post, “Mitigating Bias in A.I. Models,” citing Ms. Buolamwini’s study. In the post, Ruchir Puri, chief architect at IBM Watson, said IBM had been working for months to reduce bias in its facial recognition system. The company post included test results showing improvements, particularly in classifying the gender of darker-skinned women. Soon after, IBM released a new system that the company said had a tenfold decrease in error rates.

A few months later, Microsoft published its own post, titled “Microsoft improves facial recognition technology to perform well across all skin tones, genders.” In particular, the company said, it had significantly reduced the error rates for female and darker-skinned faces.

Ms. Buolamwini wanted to learn whether the study had changed overall industry practices. So she and a colleague, Deborah Raji, a college student who did an internship at the M.I.T. Media Lab last summer, conducted a new study.

In it, they retested the facial systems of IBM, Microsoft and Face++. They also tested the facial systems of two companies that were not included in the first study: Amazon and Kairos, a start-up in Florida.

The new study found that IBM, Microsoft and Face++ all improved their accuracy in identifying gender.

By contrast, the study reported, Amazon misclassified the gender of darker-skinned females 31 percent of the time, while Kairos had an error rate of 22.5 percent.

Melissa Doval, the chief executive of Kairos, said the company, inspired by Ms. Buolamwini’s work, released a more accurate algorithm in October.

Ms. Buolamwini said the results of her studies raised fundamental questions for society about whether facial technology should not be used in certain situations, such as job interviews, or in products, like drones or police body cameras.

Some federal lawmakers are voicing similar issues.

“Technology like Amazon’s Rekognition should be used if and only if it is imbued with American values like the right to privacy and equal protection,” said Senator Edward J. Markey, a Massachusetts Democrat who has been investigating Amazon’s facial recognition practices. “I do not think that standard is currently being met.”

A version of this article appears in print on Jan. 25, 2019, on Page B1 of the New York edition with the headline: Unmasking a Concern

