

Opinion

The Racist History Behind Facial Recognition

When will we finally learn we cannot predict people's character from their appearance?

By **Sahil Chinoy**

Mr. Chinoy is a graphics editor for The New York Times Opinion section.

July 10, 2019

Researchers recently learned that Immigration and Customs Enforcement used facial recognition on millions of driver's license photographs without the license-holders' knowledge, the latest revelation about governments employing the technology in ways that threaten civil liberties.

But the surveillance potential of facial recognition — its ability to create a “perpetual lineup” — isn't the only cause for concern. The technological frontiers being explored by questionable researchers and unscrupulous start-ups recall the discredited pseudosciences of physiognomy and phrenology, which purport to use facial structure and head shape to assess character and mental capacity.

Artificial intelligence and modern computing are giving new life and a veneer of objectivity to these debunked theories, which were once used to legitimize slavery and perpetuate Nazi race “science.” Those who wish to spread essentialist theories of racial hierarchy are paying attention. In one blog, for example, a contemporary white nationalist claimed that “physiognomy is real” and “needs to come back as a legitimate field of scientific inquiry.”

More broadly, new applications of facial recognition — not just in academic research, but also in commercial products that try to guess emotions from facial expressions — echo the same biological essentialism behind physiognomy. Apparently, we still haven't learned that faces do not contain some deeper truth about the people they belong to.

Composite photographs, new and old

One of the pioneers of 19th-century facial analysis, Francis Galton, was a prominent British eugenicist. He superimposed images of men convicted of crimes, attempting to find through “pictorial statistics” the essence of the criminal face.

Francis Galton's composite portraits of “men convicted of larceny.” • Source: Galton.org

Galton was disappointed with the results: He was unable to discern a criminal “type” from his composite photographs. This is because physiognomy is junk science — criminality is written neither in one's genes nor on one's face. He also tried to use composite portraits to determine the ideal “type” of each race, and

his research was cited by Hans F.K. Günther, a Nazi eugenicist who wrote a book that was required reading in German schools during the Third Reich.

Galton's tools and ideas have proved surprisingly durable, and modern researchers are again contemplating whether criminality can be read from one's face. In a much-contested 2016 paper, researchers at a Chinese university claimed they had trained an algorithm to distinguish criminal from noncriminal portraits, and that "lip curvature, eye inner corner distance, and the so-called nose-mouth angle" could help tell them apart. The paper includes "average faces" of criminals and noncriminals reminiscent of Galton's composite portraits.

Criminal **Noncriminal**
(composite images) (composite images)

Note: Images in the top and bottom rows were generated using different averaging techniques. • Source: Xiaolin Wu and Xi Zhang, "Automated Inference on Criminality Using Face Images" • By The New York Times

The paper echoes many of the fallacies in Galton's research: that people convicted of crimes are representative of those who commit them (the justice system exhibits profound bias), that the concept of inborn "criminality" is sound (life circumstances drastically shape one's likelihood of committing a crime) and that facial appearance is a reliable predictor of character.

It's true that humans tend to agree on what a threatening face looks like. But Alexander Todorov, a psychologist at Princeton, writes in his book "Face Value" that the relationship between a face and our sense that it is threatening (or friendly) is "between appearance and impressions, not between appearance and character." The temptation to think we can read something deeper from these visual stereotypes is misguided — but persistent.

In 2017, the Stanford professor Michal Kosinski was an author of a study claiming to have invented an A.I. "gaydar" that could, when presented with pictures of gay and straight men, determine which ones were gay with 81 percent accuracy. (He told The Guardian that facial recognition might be used in the future to predict I.Q. as well.)

Straight **Gay**
(composite images) (composite images)

Source: Michal Kosinski and Yilun Wang, "Deep neural networks are more accurate than humans at detecting sexual orientation from facial images" • By The New York Times

The paper speculates about whether differences in facial structure between gay and straight men might result from underexposure to male hormones, but neglects a simpler explanation, wrote Blaise Agüera y Arcas and Margaret Mitchell, A.I. researchers at Google, and Dr. Todorov in a Medium article. The research relied on images from dating websites. It's likely that gay and straight people present themselves differently on these sites, from hairstyle to the degree they are tanned to the angle they take their selfies, the critics said. But the paper focuses on ideas reminiscent of the discredited theory of sexual inversion, which maintains that homosexuality is an inborn "reversal" of gender characteristics — gay men with female qualities, for example.

"Using scientific language and measurement doesn't prevent a researcher from conducting flawed experiments and drawing wrong conclusions — especially when they confirm preconceptions," the critics wrote in another post.

Echoes of the past

Parallels between the modern technology and historical applications abound. A 1902 phrenology book showed how to distinguish a "genuine husband" from an "unreliable" one based on the shape of his head; today, an Israeli start-up called Faception uses machine learning to score facial images using personality types like "academic researcher," "brand promoter," "terrorist" and "pedophile."

Faception's marketing materials are almost comical in their reduction of personalities to eight stereotypes, but the company appears to have customers, indicating an interest in "legitimizing this type of A.I. system," said Clare Garvie, a facial recognition researcher at Georgetown Law.

"In some ways, they're laughable," she said. "In other ways, the very part that makes them laughable is what makes them so concerning."

In the early 20th century, Katherine M.H. Blackford advocated using physical appearance to select among job applicants. She favored analyzing photographs over interviews to reveal character, Dr. Todorov writes. Today, the company HireVue sells technology that uses A.I. to analyze videos of job applicants; the platform scores them on measures like "personal stability" and "conscientiousness and responsibility."

Cesare Lombroso, a prominent 19th-century Italian physiognomist, proposed separating children that he judged to be intellectually inferior, based on face and body measurements, from their "better-endowed companions." Today, facial recognition programs are being piloted at American universities and Chinese schools to monitor students' emotions and engagement. This is problematic for myriad reasons: Studies have shown no correlation between student engagement and actual learning, and teachers are more likely to see black students' faces as angry, bias that might creep into an automated system.

Classification and surveillance

The similarities between modern, A.I.-driven facial analysis and its earlier, analog iteration are eerie. Both, for example, originated as attempts to track criminals and security targets.

Alphonse Bertillon, a French policeman and facial analysis pioneer, wanted to identify repeat offenders. He invented the mug shot and noted specific body measurements like head length on his "Bertillon cards." With records of more than 100,000 prisoners collected between 1883 and 1893, he identified 4,564 recidivists.

Bertillon’s classification scheme was superseded by a more efficient fingerprinting system, but the basic idea — using bodily measurements to identify people in the service of an intelligence apparatus — was reborn with modern facial recognition. Progress in computer-driven facial recognition has been spurred by military investment and government competitions. (One C.I.A. director’s interest in the technology grew from a James Bond movie — he asked his staff to investigate facial recognition after seeing it used in the 1985 film “A View to Kill.”)

Early facial recognition software developed in the 1960s was like a computer-assisted version of Bertillon’s system, requiring researchers to manually identify points like the center of a subject’s eye (at a rate of about 40 images per hour). By the late 1990s, algorithms could automatically map facial features — and supercharged by computers, they could scan videos in real time.

Many of these algorithms are trained on people who did not or could not consent to their faces being used. I.B.M. took public photos from Flickr to feed facial recognition programs. The National Institute of Standards and Technology, a government agency, hosts a database of mug shots and images of people who have died. “Haunted data persists today,” said Joy Buolamwini, an M.I.T. researcher, in an email.

Emotional “intelligence”

Facial analysis services are commercially available from providers like Amazon and Microsoft. Anyone can use them at a nominal price — Amazon charges one-tenth of a cent to process a picture — to guess a person’s identity, gender, age and emotional state. Other platforms like Face++ guess race, too.

But these algorithms have documented problems with nonwhite, nonmale faces. And the idea that A.I. can detect the presence of emotions — most commonly happiness, sadness, anger, disgust and surprise — is especially fraught. Customers have used “affect recognition” for everything from measuring how people react to ads to helping children with autism develop social and emotional skills, but a report from the A.I. Now Institute argues that the technology is being “applied in unethical and irresponsible ways.”

Rekognition in Action

Amazon’s facial analysis service returns guesses about the emotions detected on a face, in addition to predictions about gender and age.

Source image

Aaron Sorkin

Output emotions

Calm	0%
Confused	2
Angry	1
Disgusted	1
Sad	0
Surprised	2
Happy	93

Output gender

Male (96% confidence)

Output age

35 to 52

Affect recognition draws from the work of Paul Ekman, a modern psychologist who argued that facial expressions are an objective way to determine someone's inner emotional state, and that there exists a limited set of basic emotional categories that are fixed across cultures. His work suggests that we can't help revealing these emotions. That theory inspired the television show "Lie to Me," about a scientist who helps law enforcement by interpreting unforthcoming suspects' expressions.

Dr. Ekman's work has been criticized by scholars who say emotions cannot be reduced to such easily interpretable — and computationally convenient — categories. Algorithms that use these simplistic categories are "likely to reproduce the errors of an outdated scientific paradigm," according to the A.I. Now report.

[If you're online — and, well, you are — chances are someone is using your information. We'll tell you what you can do about it. Sign up for our limited-run newsletter.]

Moreover, it is not hard to stretch from interpreting the results of facial analysis as "how happy this face *appears*" to the simpler but inaccurate "how happy this person *feels*" or even "how happy this person really is, despite his efforts to mask his emotions." As the A.I. Now report says, affect recognition "raises troubling ethical questions about locating the arbiter of someone's 'real' character and emotions outside of the individual."

We've been here before. Much like the 19th-century technologies of photography and composite portraits lent "objectivity" to pseudoscientific physiognomy, today, computers and artificial intelligence supposedly distance facial analysis from human judgment and prejudice. In reality, algorithms that rely on a flawed understanding of expressions and emotions can just make prejudice more difficult to spot.

In his book, Dr. Todorov discusses the German physicist Georg Christoph Lichtenberg, an 18th-century skeptic of physiognomy who thought that the practice "simply licensed our natural impulses to form impressions from appearance."

If physiognomy gained traction, "one will hang children before they have done the deeds that merit the gallows," Lichtenberg wrote, warning of a "physiognomic auto-da-fé."

As facial recognition technology develops, we would be wise to heed his words.

Sahil Chinoy is a graphics editor for The New York Times Opinion section.

Like other media companies, The Times collects data on its visitors when they read stories like this one. For more detail please see our privacy policy and our publisher's description of The Times's practices and continued steps to increase transparency and protections.

Follow @privacyproject on Twitter and The New York Times Opinion Section on Facebook and Instagram.

A version of this article appears in print on , Section SR, Page 3 of the New York edition with the headline: Facial Recognition's Racist History