EPILOGUE

"Who will ever want to be photographed by me?!" asks Adrienne, a photographer at the shared studio complex. "I don't want to be photographed myself," she adds.

She has just read journalist Tamar Stelling's article in *De Correspondent* about PimEyes, a reverse image search engine.[184] The idea of one's photographic portrait being viewed by non-human, machine spectators is indeed unsettling, and there is a good chance that sitters, anticipating such a gaze, might run from the studio. Does this mark the end of photographic portraiture as we know it? Are machine spectators yet another, and possibly the final, argument for redefining the photographic portrait? Not in pursuit of a better or more fitting portrait for the sitter and photographer but driven purely by necessity.

The interest of non-human spectators, or machines, in photographic portraits is twofold: emotion recognition and data collection. Neither scenario is particularly appealing for the sitter.

Emotion Recognition

Emotion recognition in machine vision is a subfield of artificial intelligence that focuses on teaching machines to recognize and interpret images of people. It often relies on machine learning (ML) techniques, where

184. Tamar Stelling, "Van swipe tot stalk: daten ten tijde van gezichtsherkenning," *De Correspondent*, February 14, 2024, https://decorrespondent.nl/15120/van-swipe-tot-stalk-daten-ten-tijde-van-gezichtsherkenning/b84dacd1-83f6-09a8-2235-623caa8fa9b1.

computers "learn" statistical patterns from pre-existing data sets and then use these models to identify similar patterns in new, related data.[185]

Humans have always found it useful to understand how others feel. Evolutionarily, our survival has depended on our ability to read faces and distinguish good intentions from bad ones. As previously mentioned, our brains are hardwired to do this. Many find it appealing to imagine that emotions can be extracted from facial images. If machines could do this, people's emotions may be "read" via cameras without their permission or knowledge.

The concept of machines reading people's emotions from their facial images has been warmly embraced by companies interested in understanding customer reactions to products or evaluating candidates for online job applications. Governments, too, are keen on reading emotions in public spaces (for example, to enhance security at airports). The desire to predict criminal intentions has been a major motivator for governments to advance facial recognition technology, particularly in the United States after the 9/11 attacks. However, it was not until 2015, with the number of faces online growing exponentially thanks to the popularity of Instagram and other social media platforms, that facial and emotion recognition truly began to flourish. These online faces provided the necessary data sets on which this technology relies.

What to Recognize?

Most emotion recognition systems are based on psychologist Paul Ekman's (Washington D.C., 1934) Facial Action Coding System (FACS), which stems from his

185. Ethem Alpaydin, *Introduction to Machine Learning*, 4th ed. (MIT Press, 2020).

Basic Emotion Theory (BET). This theory identifies six basic emotions – fear, anger, happiness, sadness, disgust, and surprise – along with a secondary category of "micro-expressions" that are supposedly impossible to simulate.[186]

It is tempting to believe that faces can be "read" in this way, and that distinct categories of human emotion can be universally interpreted from facial expressions. However, this is not how human emotion recognition actually works. For this reason, although FACS is widely used, it has been challenged and deconstructed by psychologists and anthropologists like emotion researcher Lisa Feldman Barrett (Toronto, 1963). After re-examining Ekman's studies, Feldman Barrett concluded that they were flawed, often based on suggestive questioning.[187] Human emotion is simply too complex to fit neatly into discrete categories. Some people laugh when they are happy, while others laugh because they are nervous. Moreover, happiness does not always translate into constant smiling. Emotions are relational and multifaceted, and it is a misconception that a face can be "read" in a split second just by deciphering an expression. Instead, people infer someone's emotional state by considering multiple factors, such as context and the events leading up to that moment.

The importance of context in recognizing human emotion is often illustrated with the example of a screaming football player in a photograph. The player is screaming, but what does the scream mean? People interpret it very differently depending on the information they are given. If told the player just scored, they see the scream as a cry of joy; if told he missed the goal, the scream becomes an expression of anger and frustration.

186. Paul Ekman and Erika L. Rosenberg, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)* (Oxford University Press, 2005).
187. Lisa Feldman Barrett, *How Emotions Are Made: The Secret Life of the Brain* (Houghton Mifflin Harcourt, 2017).

This illustrates why, according to Feldman Barrett, current emotion recognition systems fall short. People do not passively recognize emotions; they actively interpret them, relying on a variety of contextual cues such as body posture, hand gestures, words, the social setting, and the person's cultural background. This complexity is missing in current emotion recognition systems. For computers to truly understand the nuances of human emotion, they would need to observe a person over a longer period of time.

Another misreading of facial expressions occurs in the Japanese Female Facial Expression (JAFFE) database developed by Michael Lyons, Miyuki Kamachi, and Jiro Gyoba in 1998, which is widely used in affective computing research.[188] This dataset contains photographs of ten Japanese female models in a studio, making seven facial expressions that are supposed to correlate with seven basic emotional states. The purpose of the dataset is to help machine learning systems recognize and label these emotions in newly captured, unlabeled images. Ironically, these facial expressions are performed, rather than occurring naturally. They are acted out in a controlled setting, meaning that they do not necessarily reflect the internal emotional states of the models. In this case, the "reading" of people's true emotions is based on comparison with datasets of images that do not actually correspond to real emotional states.

188. M. Lyons et al., "Coding Facial Expressions with Gabor Wavelets," *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition* (1998): 200–205, https://doi.org/10.1109/AFGR.1998.670949.

Confusing Form with Meaning

The fundamental issue with datasets used for emotion or face recognition, and with artificial intelligence as a whole, lies in how the images are labeled. In the JAFFE dataset, for example, an image of a woman pretending to be happy is labeled as "happy." This label is not only

inaccurate but fundamentally wrong because the woman was not actually happy – she was pretending to be, which is entirely different.

Images do not describe themselves, and the interpretation of images – the relationship between images and meaning – is nuanced, unstable, and profoundly complex. It is a relational process. Images are elusive, laden with multiple potential meanings, unresolved questions, and contradictions. Anyone who has ever created or studied images – as an artist, art historian, philosopher, or media theorist – knows this well. Even someone simply wondering what they are seeing when looking at an image understands the complexity involved. However, as Bender points out, this critical understanding is often lost in the construction of AI training sets. These datasets conflate what something *looks like* with what it *is*.[189]

In datasets, images are labeled and categorized. At rates of up to fifty images per minute, large quantities of photographs scraped from the internet are labeled by remote workers sitting behind their computers.[190] While some labels may seem harmless at first glance, such as "happy" in the JAFFE dataset, the problem of labeling photographs becomes glaring when one tries to assign a label to a photo of a person. For instance, how does one decide whether a photographic portrait should be labeled "adventurous," "professor," or "criminal?"

Machine spectators compare new images to patterns in the training set, which consists of labeled image categories, and draw conclusions based on these comparisons. However, as AI researcher Kate Crawford (Australia, 1974) and artist Trevor Paglen (Camp Springs, 1974)

189. Bender, "Climbing towards NLU," 5186.
190. John Markoff, "Seeking a Better Way to Find Web Images," *New York Times*, November 19, 2012, https://www.nytimes.com/2012/11/20/science/for-web-images-creating-new-technology-to-seek-and-find.html.

argue, this process is built on several flawed assumptions about the nature of images, labels, categorization, and representation.[191] First, it assumes that categories such as emotions, gender, or "losers" exist as fixed and consistent concepts. Second, it assumes a universal, fixed correspondence between images and concepts, appearances and essences. It also assumes simple, self-evident, and measurable links between images, their referents, and labels. In other words, it assumes that abstract concepts – whether "happy" or "adventurous" – have some kind of visual essence, and that this essence can be identified using statistical methods to find patterns in labeled images. This means that images labeled "losers" should, in theory, contain visual patterns that distinguish them from, say, images of farmers or assistant professors.

Finally, the structure of these training sets assumes that all concrete nouns are created equally and that many abstract nouns can also be visually expressed (e.g., "happiness" or "anti-Semitism").[192]

Categories and labels attempt to impose order on a complex universe, but the impossibility of this becomes stark when we see labels applied to people. Crawford and Paglen illustrate this by searching the dataset Imagenet, one of the most widely used training sets in machine learning. They found a photograph of a child wearing sunglasses that was classified as a "failure, loser, non-starter, unsuccessful person."[193]

As Crawford and Paglen point out, these training sets are increasingly embedded in our urban, legal, logistical, and commercial infrastructures. They hold an important yet underexplored power: the ability to shape

191. Kate Crawford and Trevor Paglen, "Excavating AI: The Politics of Training Sets for Machine Learning," September 19, 2019, https://excavating.ai.
192. Crawford and Paglen, "Excavating AI."
193. Crawford and Paglen, "Excavating AI."

the world in their own image.[194] Moreover, these assumptions echo times in the past when the visual assessment and classification of people was used as a tool of oppression and racial science.[195]

Physiognomic AI

This is why media scholar Luke Stark and attorney Jevan Hutson refer to emotion recognition as "Physiognomic AI." They coined this term to describe the practice of using computer software and related systems to create hierarchies based on an individual's body composition, perceived character, abilities, and future social outcomes, all inferred from physical or behavioral characteristics. According to Stark and Hutson, the logics of physiognomy (the discredited pseudoscience of facial reading) and phrenology (the equally discredited pseudoscience of skull measurements) are deeply embedded in the technical mechanisms of computer vision applied to humans. As a result, machine learning (ML), computer vision, and related AI technologies are ushering in a new era of computational physiognomy and phrenology, reviving these outdated ideas in concept, form, and practice, and posing a threat to civil liberties.[196]

Physiognomy and phrenology rest on the premise that analyzing facial features or the skull reveals a person's "mental and physical power." Today, similar conclusions about a person's abilities or future prospects are drawn from their physical appearance or behavior. These traits can include cognitive abilities, emotional tendencies, or even the likelihood of criminal behavior. The social outcomes predicted by these systems can range from employability and creditworthiness to voting patterns and potential criminality.[197]

194. Crawford and Paglen, "Excavating AI."
195. Crawford and Paglen, "Excavating AI."
196. Luke Stark and Jevan Hutson, "Physiognomic Artificial Intelligence," *Fordham Intellectual Property, Media and Entertainment Law Journal* 32, no. 4 (2022): 922–978, https://ir.lawnet.fordham.edu/iplj/vol32/iss4/2.
197. Stark and Hutson, "Physiognomic Artificial Intelligence," 944.

However, physiognomy, and the computer vision technology based on it, is fundamentally flawed. One cannot infer a person's character or abilities simply by observing their outward appearance. This has long been recognized, and scientists across various disciplines have repeatedly demonstrated that physiognomy is an unfounded, racist, and thoroughly discredited pseudoscience.[198]

Despite the discrediting of phrenology as a scientific field and the disappearance of physiognomy from popular discourse after World War II, interest in physiognomic analysis has never entirely vanished. This is largely because physiognomic and phrenological assumptions help maintain existing racist, sexist, and classist social hierarchies.[199]

Physiognomic claims also persist due to people's tendency to "judge a book by its cover," which is deeply ingrained in our cultural habits.[200] While this human tendency is damaging on its own, the automation of this impulse through digital technologies is even more alarming.[201] Unlike in physiognomy's original heyday, these judgments are now hidden behind the labeling and categorization of images in data training sets. They are disguised by the seeming objectivity of computers. For this reason, Stark and Hutson argue that physiognomic AI is reviving scientific racism on an unprecedented scale whenever it is used to make claims about people's thoughts, preferences, or potential behavior – whether evaluating their appreciation for products, suitability for jobs, or likelihood of criminal activity.[202]

As computer scientist Arvind Narayanan (Mumbai, 1981) states in his "How to Recognize AI Snake Oil"

198. Sharrona Pearl, *About Faces: Physiognomy in Nineteenth-Century Britain* (Harvard University Press, 2010), 222.
199. Sahil Chinoy, "The Racist History Behind Facial Recognition," *New York Times*, July 10, 2019, https://www.nytimes.com/2019/07/10/opinion/facial-recognition-race.html; Catherine Stinson, "Algorithms Associating Appearance and Criminality Have a Dark Past," *Aeon*, May 15, 2020, https://aeon.co/ideas/algorithms-associating-appearance-and-criminality-have-a-dark-past.
200. Pearl, *About Faces*, 216.
201. Stark and Hutson, "Physiognomic Artificial Intelligence," 939.
202. Lisa Nakamura, *Cybertypes: Race, Ethnicity, and Identity on the Internet* (Routledge, 2002).

presentation, AI's ability to predict such social outcomes is fundamentally questionable.[203] In Narayanan's words, "We can't predict the future. That should be common sense. But we seem to have decided to suspend common sense when AI is involved."[204]

It's a troubling scenario for sitters to have their photographic portraits scrutinized by machine spectators searching for emotions. Even if the camera in a studio is not connected to software that "reads" emotions and makes superficial, misleading claims about the subject, there is still a significant risk that the portrait could unintentionally end up in a database – perhaps via the photographer's website or social media – where it may be scraped and added to an image database. From there, it could contribute to pseudoscientific physiognomic AI.

(Un)interested Machines

Machine spectators also examine photographic portraits to gather data. In this context, a portrait functions as a key to other images and online information about the person. Through reverse image searches, the portrait is scanned to link databases containing the same face, connecting digital traces of the individual – such as holiday photos, traffic violation snapshots, or social media images where the sitter might appear in the background.

It's difficult to fully grasp the implications of a world without privacy, where walking down the street anonymously has vanished. In China, for instance, nearly one billion "smart cameras" are connected to facial recognition systems linked to "social credit," where even minor infractions – like ignoring a red light – can have conse-

203. Arvind Narayanan, "How to Recognize AI Snake Oil," Arthur Miller Lecture on Science and Ethics, Massachusetts Institute of Technology, November 18, 2019, https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf..
204. Narayanan, "How to Recognize AI

quences, such as difficulty in applying for a mortgage. Similarly, in the Netherlands, there are an estimated 1.2 million cameras illegally monitoring the streets, capable of recording and sometimes analyzing everyone who passes by.[205] What happens when one's past is always publicly accessible, both on a personal level and as a society? How does an adolescent develop their identity when there is no space to leave behind what they no longer want to be? How does change happen when (totalitarian) regimes can control any possible dissonance? As tech philosopher Evgeny Morozov (Soligorsk, 1984) suggests, what if Rosa Parks had never boarded the bus because the bus door wouldn't open for a Black face?[206]

What about everyday life? Strangers in a bar could quickly snap a photo of you and instantly find all your information online, including your address. Glasses equipped with reverse image search technology might soon make even taking a photo unnecessary – simply pointing the glasses at someone could project all the images and data retrieved online about that person onto the lens.

Algorithms that link a face in a photo to other online images, essentially a "Google for faces", have been in development since 2016, including by the founder of Clearview AI. In her book *Your Face Belongs to Us*, *New York Times* tech reporter Kashmir Hill (US, 1981) describes how Clearview AI goes beyond other companies by scraping millions of photos from social media sites such as Facebook, Twitter, LinkedIn, and Instagram. These photos include not only people posing but also bystanders accidentally captured in the background.[206] Clearview AI's app licenses have been sold to

205. Iva Venneman and Pieter Sabel, "Slimme deurbel leidt tot hausse and privacyklachten," *de Volkskrant*, February 23, 2024.
206. Alexis C. Madrigal, "Toward a Complex, Realistic, and Moral Tech Criticism," *The Atlantic*, March 13, 2013, https://www.theatlantic.com/technology/archive/2013/03/toward-a-complex-realistic-and-moral-tech-criticism/273996/.
207. Kashmir Hill, *Your Face Belongs to Us: The Secretive Startup Dismantling Your Privacy* (Simon & Schuster, 2023).

several U.S. police departments, who use it to find individuals resembling photographic images of criminals. Unsurprisingly, there have been cases where innocent people were stopped or even arrested simply because their photo appeared in Clearview's database.

What is particularly frightening about Hill's book is that this small start-up was able to gather all the information it needed from freely available online sources - and managed to create the most powerful facial recognition search engine to date. Moreover, the book reveals the immense power that can be wielded by individuals driven by technological progress but unencumbered by ethical concerns or consideration of the consequences.

These two scenarios are not very appealing for the sitter. In the physiognomic AI emotion recognition variant, the photographic portrait may unwillingly become part of data training sets used to make judgments about people. In the second scenario, the sitter's portrait becomes part of a web of information surrounding them, with every image of the sitter online contributing to an increasingly tighter web, making it harder to present oneself differently from what is already visible in the past.

The article Adrienne read in *De Correspondent* about PimEyes explains the reverse image search engine, which works similarly to Clearview AI.[208] Like Clearview AI, but available to anyone for €35 a month, PimEyes allows users to enter a photographic portrait (or hold their iPhone in front of someone's face), and the site will return numerous photos of that person from various websites. While this might be convenient for finding information that Tinder dates did not share in

208. Stelling, "Van swipe tot stalk."

their profiles, it has far broader, potentially invasive applications – such as identifying people at a demonstration or uncovering the hidden pasts of colleagues.

Although it may be fun to find information about others, it is far less comfortable to imagine what can be discovered about oneself. As a result, Stelling predicts that people will likely start adopting new ways of handling their photographic images. Schools in Amsterdam, for example, have stopped taking class photos for fear of GDPR-related claims, and clubs have begun taping over smartphone lenses to provide a safe space where no pictures are taken.

This is exactly what Adrienne fears: Who would willingly sit in front of a camera knowing that their photograph could become part of an ever-tightening web of visual information?

"Well, there are some tricks, I think," I say, trying to reassure Adrienne. "I briefly skimmed some information online, and it seems there's something about removing certain pixels to make images unrecognizable by machines."

"I want that!" Adrienne eagerly responds. "Please send it to me if you find anything."

That evening, as I search for "data poisoning 2024," it becomes clear that I will probably have to disappoint her. While there were hopeful developments like Fawkes and LowKey, tools designed to use adversarial machine learning techniques to disrupt images before they are posted online, they no longer seem effective. The idea behind these tools was to poison the facial recognition

models trained on these images. Unfortunately, I quickly come across an article explaining that these strategies do not work and merely provide a false sense of security. The authors suggest we place our hopes on legislation instead. Since 2022, there has been an eerie silence around potential countermeasures.[209]

Many artists have tackled the issue of facial recognition to raise awareness, such as the Dazzle Club, a group of art students who, in 2021, marched through the streets of London wearing geometric face paint to "dazzle" facial recognition systems. I have also noticed more and more profile photos of people with their eyes closed or with ping-pong balls over their eyes – presumably to confuse the algorithm.

I briefly fantasize about developing a counter-practice with photographic portraits, something that would poison the data training sets and resist this development. But I quickly realize I do not know how. Perhaps this is the moment to think differently about photographic portraiture, as Stelling predicts. Maybe this really is the time to rethink how we see photographic images. This could actually be the end of the photographic portrait as we know it.

Perhaps now *is* the moment to embrace situative portraits. Only this time, the need for situative portraits aligns not just with our contemporary understanding of identity and who we think we are but also with the kind of social environment in which we want to live.

209. Evani Radiya-Dixit, Nicholas Carlini, Sanghyun Hong, and Florian Tramèr, "Data Poisoning Won't Save You from Facial Recognition," paper presented at the International Conference on Learning Representations (ICLR), 2022, https://openreview.net/pdf?id=B5XahNLmna.