MEDIA FIELDS
J O U R N A L

Gender Estimation in Face Recognition Technology: How Smart Algorithms Learn to Discriminate

Sarah Kember

Face recognition technology is becoming central to a naturalized—embedded and invisible—ontology of everyday control. As a marketing and surveillance-based biometric and photographic technology, one of its main advantages over other biometrics such as finger printing or iris-scanning is that it operates at a distance and does not require consent or participation. Face recognition is a default setting on social networking sites like Facebook, offering automatic tagging suggestions as the user uploads photographs of friends and family. It is becoming ubiquitous in international airports and other social environments where security and/or commerce are at stake. The objective of face recognition is to be able to pick out a face from a crowd and identify the target by comparing it with a database. Where this objective is hard to achieve, another goal entails learning to discriminate between classes of faces based on gender, race, and age. This is easier in that it relates to groups rather than individuals and appeals to biological differences.

Face recognition systems seek to overcome the division between human and machine vision or, specifically, between human and machinic capacities for appearance-based face recognition and identification. Questions of system accuracy and performance come to the fore because the comparison remains unfavorable. Subsequent performance anxiety serves to legitimize a range of technological innovations designed to close the gap, and among them is the use of AI: "AI approaches utilize tools such as neural networks and machine learning techniques to recognize faces." How do machines learn? The issue has been widely debated but in this context it is clear that in addition to techniques of pattern recognition and sorting, the principle mechanism of machine learning is reductionism. Matthew Turk and Alex Pentland have made a significant contribution to the development of face recognition. For them, "developing a computational model of face recognition is quite difficult, because faces are complex, multidimensional, and meaningful visual stimuli."[1] Face recognition systems substitute the meaning of faces for a mathematics of faces, reducing their complexity and multidimensionality to measurable, predictable criteria. Moreover, face recognition technology requires a reduction in the variation of face images and environments and must ultimately replace faces with vectors (principal components of faces) or with standardized templates in order to learn anything at all. System accuracy and performance depend on "constrained environments such as an office or a household."[2] The face image presented to the system for recognition must be centered, "the same size as the training images," and fully frontal or in profile, so reproducing—as input—the mug shot photograph generated by nineteenth-century ways of seeing.[3] An elision of labor secures the illusion of autonomy in face recognition technology. There is also an inventory of technological failures that, combined with reductionism, delimit the claim to smartness implicit in the system's ability to learn.[4] However, we cannot simply dismiss the claim, as it is manifested in the very architecture of the system. Here, smartness materializes in pattern-recognizing and sorting algorithms that are learning to identify faces by discriminating among them, generating ontological and epistemological divisions—between male and female, black and white, old and young—that in this case must remain un-reconciled, reduced to a set of essentialized categories that guarantee system performance by ensuring that input (a recognizable face) is equivalent to output (a recognized face).

The aim of a facial recognition system is either to verify or to identify someone from a still or video image. Following the acquisition of this

"probe" image, the system must first of all detect the face or distinguish between the face and its surroundings. To do this it selects certain landmark features in order to compare them with the database. Either that or it generates what are called standard feature templates—averages or types. Once detected, the face is normalized, or rather the image is standardized with respect to established photographic codes such as lighting, format, pose, and resolution. Again, this aids comparison with the database. However, the normalization algorithm is only capable of compensating for slight variations, so the probe image must already be as close as possible to a standardized portrait. In order to facilitate face recognition, the already standardized image is translated and transformed into a simplified mathematical representation called a biometric template. The trick, in this process of reductive computation, is to retain enough information to distinguish one template from another and thereby reduce the risk of creating "biometric doubles."[5]

One of the algorithms used in face recognition is Principal Component Analysis (PCA). It produces images akin to Francis Galton's nineteenth-century eugenicist photographic composites by removing extraneous information, including the outline of the face itself.[6] PCA reduces faces to their vectors and refigures them as eigenfaces. In "Eigenfaces for Recognition," Turk and Pentland explain that the system functions "by projecting face images onto a feature space that spans the significant variations among known face images."[7] Significant features are referred to as eigenfaces "because they are the eigenvectors (principal components) of the set of faces." They may correspond to familiar features like eyes and noses whose geometric relation is then measured and computed. Each input, or individual face image, is "a weighted sum of the eigenface features, and so to recognize a particular face it is necessary only to compare these weights to those of known individuals." Turk and Pentland acknowledge that an eigenface is "an extremely compact representation" not only of the face but of the original face image.[8] It is a practical rather than elegant solution to the problem of face recognition.[9]

Linear Discriminant Analysis (LDA) is another key algorithm. It creates classes of faces, much like Havelock Ellis did in his nineteenth-century physiognomy of criminals. In their survey of face recognition techniques, Jafri and Arabia explain that LDA "maximises the ratio of the between-class scatter" and is better at classifying and discriminating between classes of faces than PCA.[10] This may be partly because this approach starts by selecting faces that are already distinctive. As LDA researchers Kamran Etemad and Rama Chellappa state, "First, we need a training set

composed of a relatively large group of subjects with diverse facial characteristics. The appropriate selection of the training set directly determines the validity of the final result."[11] Sorting algorithms discriminate between classes and types of faces. Both LDA and, increasingly, PCA are being used to discriminate on the basis of gender.

Contemporary face recognition systems differ from earlier analogue and digital systems in that they are exclusively oriented toward recognition rather than recall. They are designed according to the surveillance and marketing imperatives of targeting, tracking, and location. However, picking out one face in a crowd is harder and more prone to error than identifying once class of faces as distinct from another, especially when that class appeals to the biological categories that inform gender, race, and age. These categories are naturalized through geometric coding techniques (where syntactic coding is reserved for face recall) and the default subject of these techniques is still the young white male.

Jacque Penry's PhotoFIT pack came in to use in the 1970s and consisted of photographic images of five features (hair and forehead, eyes, nose, mouth, and chin) mounted on card.[12] He included a male and female database but established what he claimed was a universal—genderless—facial topography. This was actually derived from a norm, a young white male that face recognition systems continue to use, but with the aim, for example, of "restricting access to certain areas based on gender" or "collecting valuable demographics" such as "the number of women entering a retail store on a given day."[13] The segue from disciplinary to biopower is, for Foucault, contingent on the increasing use of demographics and statistics that orient governance more towards the populace than the individual.[14] Face recognition systems demonstrate both forms of power and perhaps even the shift from one to the other. This becomes clearer as we track back from the biopolitical uses and applications of face recognition technology to the disciplinary design and architecture of the technology itself.

Koray Balci and Volkan Atalay present two algorithms for "gender estimation."[15] They point out that the same algorithms can be used "for different face specific tasks" such as race or age estimation, "without any modification."[16] In the first algorithm, the training face images are normalized and the eigenfaces are established using PCA.[17] PCA is described here as a statistical technique for "dimensionality reduction and feature extraction."[18] The performance of the system is improved by the subsequent use of a "pruning" algorithm, which identifies statistical

connections extraneous to gender (race or age) estimation and deletes them. "After deletion, the system is re-trained" and the pruning is repeated until "all the connections are deleted."[19] A performance table is produced, showing the relation between each iteration of pruning, the percentage of deleted connections, and the accuracy of the system. The accuracy of gender estimation in Balci and Volkan's experiment actually diminishes after the eighth iteration, albeit by only a few percentage points, allowing them to claim that the system is stable. They maintain that pruning or the deletion of statistical connections improves gender estimation not in a linear or absolute sense but by enhancing the process of classification itself.

For Geoffrey Bowker and Susan Leigh Star, classification is a largely invisible, increasingly technological, and fundamentally infrastructural means of "sorting things out."[20] It is an instrument of power-knowledge that is productive of the things it sorts—things such as faces that are by no means "unambiguous entities" that precede their sorting.[21] The existence of a pruning algorithm that renders faces less ambiguous testifies to their elusiveness, or their inherent resistance to classification as one mode of representationalism. It would, perhaps, be going too far to suggest that there is a crisis of representationalism in appearance-based face recognition systems. However, their designers and engineers are clearly aware that faces are things that "resist depiction"[22] because they are "complex and multidimensional"[23] and not "unique, rigid" objects.[24] The advantage of a more dynamic and relational approach to the production of faces in face recognition technology would include recognizing representationalism as a claim, a defensive manoeuvre in the face of faces' non-essential ontology and dynamic co-evolution with technological systems. Still, this defensive manoeuvre matters in a double sense: it is both meaningful and material, reproducing norms—for example, norms of gender in a machine that is learning to classify, sort, and discriminate among the population—better than it could before. If this is a last push to representationalism, it is one that reinforces it rather than shows it the door. Face recognition technology upholds a belief in the existence of ontological gaps between representations and that which they represent. It also re-produces the norms of nineteenth-century disciplinary photography even as photography becomes allied to the security-based biopolitics of computational vision and smart algorithmic sorting. In this sense, Kelly Gates is right to argue that new vantage points can underscore old visions as well as old claims to unmediated visuality.[25] Like her, I question the autonomy of face recognition systems without denying that, in conjunction with human input of various kinds, they enact

what Barad calls "agential realism," generating both categories and entities by cutting and sorting male from female, black from white, old from young.[26] In a context in which security systems are fully integrated with those of marketing, these particular epistem-ontologies intersect in predictable ways with the category of criminal/citizen-consumer.[27] Since the events of 9/11, the stereotypical face of terror (gendered, racialized) has been perhaps the most represented and most elusive of all. If the problem, from a system point of view, is that the categories leak and the classification structure does not hold, the solution is to reinforce it by pruning it. This process of agential cutting and sorting strengthens statistical groups by deleting connections between them and is precisely the point of a possible intervention, the means by which the biopolitics and ethics of computational vision can be intercepted in order to make a difference.

What are the opportunities for intervention and revision in face recognition technology? I have already signalled the operational failures and technological limitations by means of which the system deconstructs. Face recognition fails in uncontrolled environments. It cannot cope with poor lighting or resolution, struggles with facial hair and glasses, and can only sort six basic types of expression, which it must produce by reducing variation and automating expression analysis.[28] In addition, face recognition technology remains over-reliant on inputting the frontal flat-lit mug shot—further trimmed to remove hair and face outline—in order to generate gendered stereotypes and generic differences. Categories of male/female, black/white, old/young are pruned at the boundary and connections are deleted. But what if they weren't? What if a sorting algorithm became a connecting algorithm by means of the substitution of a few basic instructions:

> **for all** connections **do not** compute error gradient
>
> **end for**
>
> compute threshold for connection
>
> add connections according to threshold
>
> **until** all connections are completed[29]

Software writing could not by itself retell the story of face recognition, but it might, as Anderson and Pold suggest, be a good place to start.[30] The opportunity is clearly presented by the fact that the system struggles with

ambiguity, including, especially, gender ambiguity. For one research lab, the line between male and female is neither straight nor certain. Moghaddam and Yang draw it as a curvy, snaking, incomplete trajectory with faces on either side but very close to the boundary. "It is interesting", they write, "to note not only the visual similarity of a given pair but also their androgynous appearance."[31] Indeed it is, especially when furnished with the additional insight that there are "higher error rates in classifying females" which is "most likely due to the general lack of prominent and distinct facial features in these faces."[32] If, as Brunelli and Poggio report, "eyebrow information" is considered to be amongst the most helpful in discriminating male from female, then those who wish to embrace their androgyny know not to pluck.[33]

A connecting algorithm would re-cognize, re-think faces as female-male-black-white-old-young. These faces would constitute feminist, anti-racist, anti-ageist figurations, performative images and political imaginaries akin to the cyborg. They would make manifest a non-discriminatory politics and ethics predicated on entanglement and relationality if not—or not yet—symmetry. While it is not a solution to the problem of asymmetric power relations, relationality is a means of acknowledging, a good start in taking responsibility for the fact that "what is on the other side of the agential cut is not separate from us."[34] A connecting algorithm would take a leaky boundary and play with it in order to envision the world of faces with more potential for ambiguity. This potential is not limitless—the algorithm completes its connections—because classificatory cutting and sorting is "human"; it is what human-machine assemblages do.

Notes

1 Matthew Turk and Alex Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience* 3, no. 1 (1991): 71.

2 Ibid.

3 Ibid., 76.

4 See Lucas Introna and Helen Nissenbaum, *Facial Recognition Technology: A Survey of Policy and Implementation Issues* (New York University: The Center for Disaster Preparedness and Response, August 2009), http://www.nyu.edu/ccpr/pubs/Niss_04.08.09.pdf.

5 Ibid.

6 Turk and Pentland, "Eigenfaces," 71.

7 Ibid.

8 Ibid., 73.

9 Ibid., 84.

10 Jafri and Arabnia, "Face Recognition," 47.

11 Kamran Etemad and Rama Chellappa, "Discriminant Analysis for Recognition of Human Face Images," *Journal of the Optical Society of America* 14, no. 8 (1997): 1726.

12 Sarah Kember, *Virtual Anxiety. Photography, New Technologies and Subjectivity* (Manchester: Manchester University Press, 1998).

13 Baback Moghaddam and Ming-Hsuan Yang, "Learning Gender with Support Faces," IEEE 24 (2002): 711.

14 Michel Foucault, *The Birth of Biopolitics: Lectures at the Collège de France, 1978–79* (New York: Palgrave Macmillan, 2008).

15 Koray Balci and V. Atalay, "PCA for Gender Estimation: Which Eigenvectors Contribute?" IEEE 3 (2002): 363–366.

16 Ibid., 364.

17 With Multi-Layer Perceptron (MLP) gender classifier.

18 Ibid.

19 Ibid., 365.

20 Geoffrey C. Bowker and Susan Leigh Star, *Sorting Things Out. Classification and Its Consequences* (Cambridge, MA: The MIT Press, 2002).

21 Ibid., 320.

22 J. Elkins, *Six Stories from the End of Representation* (Stanford, California: Stanford University Press, 2008): xv.

23 Turk and Pentland, "Eigenfaces," 71.

24 Jafri and Arabnia, "Face Recognition," 42.

25 Kelly Gates, *Our Biometric Future: Facial Recognition Technology and the Culture of Surveillance* (New York: New York University Press, 2011).

26 Karen Barad, *Meeting the Universe Halfway* (Durham, NC: Duke University Press, 2007).

27 David Lyon, *Surveillance after September 11* (Cambridge, UK: Polity, 2008).

28 Caifeng Shan and Ralph Braspenning, "Recognizing Facial Expressions Automatically from Video," in *Handbook of Ambient Intelligence and Smart*

29 This is a rewriting of Balci and Atalay's pruning algorithm (365).

30 Cristian Ulrik Anderson and Søren Pold, "The Scripted Spaces of Urban Ubiquitous Computing: The experience, poetics, and politics of public scripted space," *Fibreculture Journal* 19 (2011), http://nineteen.fibreculturejournal.org.

31 Moghaddam and Yang, "Learning Gender," 710.

32 Ibid. 710.

33 Roberto Brunelli and Tomaso Poggio, "HyberBF Networks for Gender Classification," *DARPA Image Understanding Workshop Proceedings* (1995), 311–314.

34 Barad, *Meeting the Universe*, 393.

Sarah Kember is a writer and Professor of New Technologies of Communication at Goldsmiths, University of London. Her work incorporates new media, photography, and feminist cultural approaches to science and technology. Publications include a novel and a short story, *The Optical Effects of Lightning* (Wild Wolf Publishing, 2011) and "The Mysterious Case of Mr Charles D. Levy" (Ether Books, 2010). Experimental work includes an edited open access electronic book entitled *Astrobiology and the Search for Life on Mars* (Open Humanities Press, 2011) and "Media, Mars and Metamorphosis" (*Culture Machine*, vol. 11). Her latest monograph, with Joanna Zylinska, is *Life After New Media: Mediation as a Vital Process* (MIT Press, 2012). She co-edits the journals of *Photographies* and *Feminist Theory*. Previous publications include:*Virtual Anxiety: Photography, New Technologies and Subjectivity* (Manchester University Press, 1998); *Cyberfeminism and Artificial Life* (Routledge, 2003) and the co-edited volume *Inventive Life: Towards the New Vitalism* (Sage, 2006). Current research includes a funded project on digital publishing and a feminist critique of smart media.