FACIAL RECOGNITION AND AI/ML TECHNOLOGIES

Reed Ballesteros

Northwestern University, Introduction to Data Science

October 31, 2021

AI/ML Focus

Facial recognition (FR) is the computing ability to identify faces in a picture. The integration of Artificial Intelligence (AI) and Machine Learning (ML) into FR technology has led to a wide, mainstream adoption in both commercial and consumer use. FR technology was introduced as early as the 1960s though the work of Woodrow W. Bledsoe, Charles Bisson, and Helen Chan developing programs for the FBI (Nilsson 2019). More advances were made in the field in the 1970s, such as FR automation by Takeo Kanade (Kanade 2016, 9:15). Despite these advancements, FR was still in its infancy and imaging input was limited to low-resolution pictures requiring a front-facing pose, a well-lit environment, and a standard expression (Nilsson 2019). The 1990s brought about transitioning from a geometric approach to a more statistical approach with template matching and the development of eigenfaces.

The late 1990s introduced deep-learning AI and ML methods with neural networks, gaining significant performance improvement (Lin et al. 1997). Increased computing power with multi-core CPUs and GPUs, improved digital imaging, and the availability of extremely large image datasets, all combined with evolving deep-learning convolutional neural networks (CNNs) has greatly improved FR performance in both accuracy and speed in recent years, as well as being able to identify faces without constraint, such as different poses and varying light sources and environments. With such advancements FR has been integrated into many more everyday use cases, the most common being able to unlock mobile phones, approve transactions or payments on the device, and install and access apps.

While FR is continuing to be used in more products and services, the technology is still far from perfect. MIT computer scientist Joy Buolamwini described in the documentary “Coded Bias” about the FR system she implemented into a college project did not recognize her face. Buolamwini, a black woman, found that the system works when she wore a white mask (Kantayya 2020, 1:45). This experience led her to collaborate with fellow computer scientist Timnit Gebru on the “Gender Shades” project which investigated commercial FR technologies

from Microsoft, IBM, and Face++ (Buolamwini et al. 2018). They found these FR systems to be biased against people of color and women, particularly black women. The datasets used to train these FR algorithms were skewed towards lighter-skinned men, which in turn does not work as well with darker-skinned women (Kantayya 2020, 9:36).

Despite these failings, FR technology has also been widely adopted by law enforcement (LE) agencies around the world for surveillance. If the technology cannot properly recognize Buolamwini's face, what effect would its shortcomings have on people of color and their communities if used by LE? Unchecked misidentification from a FR system can lead to a flawed feedback loop of more stops and arrests to a marginalized community already in dire straits.

In recent years the Chinese government has made a push to become a leader in AI and ML (Fanning et al. 2019). With that push they too, like LE, has also adopted FR technology to monitor its citizens. Government projects such as 'Sharp Eyes' and 'SkyNet' were implemented to create a network of 200 million cameras installed across the country (Gershgorn 2021). With this network, for example, jaywalkers can be detected in real-time and posted on billboards with their names to publicly shame them and warn others not to break the law (Davies 2021). Comedian John Oliver describes the Chinese government's video surveillance as the "Eye of Sauron but instead of scouring Middle Earth for the One Ring, he was just really into knowing where all his orcs like to go to dinner" (Oliver 2020). During the Hong Kong protests in 2019 demonstrators spray-painted surveillance cameras, wore masks, and used umbrellas to obscure themselves while the arresting police hid their badges to prevent their identities from being exposed (Mozur 2019).

AI/ML Method Drill-Down

Modern FR models today use some type of deep convolutional neural network (CNN) with supervised learning. A CNN starts with an input layer of features. With an image used in FR, each feature is a pixel of the image. Each feature from the input layer is fed into specific neural nodes of the following inner convolutional layer. Each node in the convolutional layer has an assigned weight for each of their inputs, adds them all together and transforms it though an activation function before its output is passed on to specific nodes of the next convolutional layer. There could be a few of these inner layers before the process reaches the final output layer and the result. While learning, the CNN will calculate its error from the desired output and adjusts the weights of each input of node in the network by "backpropagating" to the input layer and doing it again with a newly adjusted set of weights and a new image to process.

One of the most influential deep CNN framework architectures used in modern FR models is AlexNet. Developed by computer scientist Alex Krizhevsky, the architecture consists of eight layers, a mix of five fully connected layers (in which the inputs from the layer before are fed into all nodes of the current layer) and three convolutional layers. Two GPUs and are used in parallel to help with processing layers. While CNNs traditionally use hyperbolic tangent or sigmoid functions as the activation function, AlexNet used the Rectified Linear Unit (ReLU) nonlinearity function which considerably improved performance with its simple equation *max(0,x)*. While the model is in learning mode the "dropout" regularization technique is applied where random neurons of the CNN are turned off in backpropagation which in turn improved model overfitting (Krizhevsky et al. 2017).

AlexNet started in Krizhevsky's bedroom at his parents' house on a bet, resulting in a massive deep CNN consisting of 650,000 neurons and 60 million weights. The original model was submitted to the 2012 ImageNet competition where it "obliterated" the competition (Christian 2020). Since then, FR development has switched to deep CNNs where companies such as Microsoft and Google have embraced the architecture (Krizhevsky et al. 2017).

Reflection

IBM acknowledged Buolamwini and Gebru's work which led them to improve the representation in the datasets used in their Watson Visual Recognition software (Puri 2018). In the wake of the George Floyd killing and protests in the summer of 2020, IBM announced its discontinuation. Microsoft announced they would not sell their FR technology to LE. Amazon announced a moratorium for selling their Rekognition FR software for police use (Allyn 2020). Congress introduced two bills, one that limits the limits the use of FR technologies by federal and state governments (Congress.gov 2020); and another that could prevent LE agencies from using Clearview AI, a popular FR system which allegedly uses image data "illegally obtained" from social media (Robertson 2021). While this is a step in the right direction, there is still more work to do. The Chinese government continues to broaden its biometric surveillance network of cameras along with their "social credit" score to maintain order and rule over its citizens.

CNNs are used in many AI/ML-based technologies, not just in FR. Biases in FR systems are similar to ones found in other automated decision-making processes. These biases can have an effect in the human resources field such as automated systems that scan resumes or evaluate employees. Biases can influence loan or college application screening software. If a system's training datasets are skewed, they too will reflect that skewness as well. It would be a problematic feedback loop of continuing systemic issues if we deploy these faulty systems for use in high-stakes situations at scale. While FR technology has made technology more efficient, convenient, and secure, there is also a cost. Author Kai Strittmatter says "It doesn't even matter whether [FR technology is] true or not, as long as people believe it. Once you believe it's true, it's like you don't even need the policemen at the corner anymore, because you're becoming your own policeman" (Davies 2021). We need to be aware of the potential biases that can occur in any AI/ML-based system and make sure the training sets we use and the algorithms we develop fairly represents the communities they will be deployed upon. It is our responsibility as data scientists to do things the right way and understand the broad consequences which may occur.

Annotated Bibliography

Buolamwini, Joy, and Timnit Gebru. 2018. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.” *Proceedings of Machine Learning Research - Conference on Fairness, Accountability, and Transparency* 81: 77-91.

This research paper reviews three commercial gender classification systems (Microsoft, IBM, Face++) on their performance accuracy, particularly on darker-skinned females. Buolamwini and Gebru initially found standard testing datasets to be skewed towards lighter-skinned people and introduced a more balanced dataset. From their tests they found substantial disparities in accuracy classifying gender between groups of skin tone and gender. The website for their study can be found at http://gendershades.org.

Gershgorn, Dave. 2021. "China’s ‘Sharp Eyes’ Program Aims to Surveil 100% of Public Space." *Medium*, March 2, 2021. Retrieved 2021-10-30. https://onezero.medium.com/chinas-sharp-eyes-program-aims-to-surveil-100-of-public-space-ddc22d63e015

This article discusses the various camera-based surveillance projects led by several Chinese government agencies, some dating as far back as 2003. These projects help build a network of around 200 million cameras across China at the time of this writing. The article also touches upon FR technology used against the Uighur ethnic minority.

Kanade, Takeo. 2015. "Takeo Kanade: Computer Face Recognition in its Beginning." YouTube. Video, 25:10. Accessed 2021-10-29. MITCBMM. https://www.youtube.com/watch?v=fY98kQWxJQc

Computer scientist Takeo Kanade presents a comprehensive look back into early FR technology in the 1960s and 1970s, while adding insights to his own work.

Kantayya, Shalini, director. 2020. *Coded Bias*. 7th Empire Media.

This documentary follows computer scientist Joy Buolamwini and her studies and investigations into algorithmic bias based on her early experiences with FR technology. The film includes insights from experts in the field of AI and ML including Meredith

Broussard, Cathy O'Neil, Amy Webb, and Zeynep Tufekci. Includes Buolamwini's visit to Washington D.C. for her testimony on bias to the U.S. House of Representatives.

Oliver, John. 2020. *Last Week Tonight with John Oliver*. Season 7, Episode 15. "Facial Recognition." Aired June 14, 2020, on HBO.

John Oliver presents a comedic take on FR technology in his comedy-news show, which uses some sources and events acknowledged in this essay.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey Hinton. 2017. "ImageNet Classification with Deep Convolutional Neural Networks." *Communications of the ACM* 60, no. 6: 84–90. https://doi.org/10.1145/3065386.

This research paper describes the architecture behind the deep CNN AlexNet created by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. It provides an introduction, describes the dataset used, details the architecture, reports their results, and gives their conclusions. The original model was submitted to the 2012 ImageNet competition where it won, and influences FR development today.

References

Buolamwini, Joy, and Timnit Gebru. 2018. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.” *Proceedings of Machine Learning Research - Conference on Fairness, Accountability, and Transparency* 81: 77-91.

Christian, Brian. 2020. *The Alignment Problem: Machine Learning and Human Values.* New York: W.W. Norton & Company.

Congress.gov. 2020. "S.4084 - 116th Congress (2019-2020): Facial Recognition and Biometric Technology Moratorium Act of 2020." June 25, 2020. https://www.congress.gov/bill/116th-congress/senate-bill/4084.

Davies, Dave. 2021. "Facial Recognition and Beyond: Journalist Ventures Inside China's 'Surveillance State.'" *NPR*, January 5, 2021. Retrieved 2021-10-30. https://www.npr.org/2021/01/05/953515627/facial-recognition-and-beyond-journalist-ventures-inside-chinas-surveillance-sta

Fanning, David, and Neil Docherty, producers. 2019. *Frontline*. Season 2019, Episode 5. "In the Age of AI." Aired November 5, 2019, on PBS.

Gershgorn, Dave. 2021. "China’s ‘Sharp Eyes’ Program Aims to Surveil 100% of Public Space." *Medium*, March 2, 2021. Retrieved 2021-10-30. https://onezero.medium.com/chinas-sharp-eyes-program-aims-to-surveil-100-of-public-space-ddc22d63e015

Lin, Shang-Hung, Sun-Yuan Kung, and Long-Ji Lin. 1997. "Face Recognition/Detection by Probabilistic Decision-Based Neural Network." *IEEE Transactions on Neural Networks* 8, no 1: 114-132.

Kanade, Takeo. 2015. "Takeo Kanade: Computer Face Recognition in its Beginning." YouTube. Video, 25:10. MITCBMM. https://www.youtube.com/watch?v=fY98kQWxJQc

Kantayya, Shalini, director. 2020. *Coded Bias*. 7th Empire Media.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey Hinton. 2017. “ImageNet Classification with Deep Convolutional Neural Networks.” *Communications of the ACM* 60, no. 6: 84–90. https://doi.org/10.1145/3065386.

Mozur, Paul. 2019. "In Hong Kong Protests, Faces Become Weapons." *The New York Times*, July 26, 2019. Retrieved October 30, 2021. https://www.nytimes.com/2019/07/26/technology/hong-kong-protests-facial-recognition-surveillance.html

Nilsson, Nils J. 2019. *The Quest for Artificial Intelligence: A History of Ideas and Achievements.* Cambridge, MA: Cambridge University Press: 172-173.

Oliver, John. 2020. *Last Week Tonight with John Oliver*. Season 7, Episode 15. "Facial Recognition." Aired June 14, 2020, on HBO.

Puri, Ruchir. 2018. "Mitigating Bias in AI Models." *IBM Research Blog (blog), IBM*. February 6, 2018. Retrieved 2021-10-30. https://www.ibm.com/blogs/research/2018/02/mitigating-

bias-ai-models/

Robertson, Adi. 2021. "Lawmakers propose ban on police buying access to Clearview AI and other data brokers" *The Verge*, April 21, 2021. Retrieved 2021-10-30. https://www.theverge.com/2021/4/21/22395650/wyden-paul-fourth-amendment-is-not-for-sale-act-privacy-data-brokers-clearview-ai