Faces Reveal More Than Expression

Francis Scott

Binghamton University

Abstract

The rise of facial recognition technology sparks a desire for researchers to discover a method that maximizes recognition accuracy. This paper outlines the basic architecture of a convolution neural network (CNN), a deep learning approach, and the computations involved in linear discriminant analysis (LDA), a non-deep learning approach, for facial recognition purposes. CNNs have shown ability to outperform the LDA method in regards to rate of recognition. If CNNs pose an improvement to facial recognition technology, people scanned with facial recognition technology will fall victim to privacy intrusion they may be unaware of. An image of an individual’s face can potentially link to a reveal of private information. The abundance of surveillance systems and law enforcement agencies having easy access to facial recognition technology puts the public at risk if regulations are not correctly followed. Facial recognition software has the potential for abuse if in the wrong hands. The potential privacy risks surrounding facial recognition technology are explored in this paper to shed light on the impact facial recognition technology can have on society.

Faces Reveal More Than Expression

**Introduction**

Biometric facial recognition systems are becoming more prevalent, and the accessibility of these systems pose alarming privacy concerns. A person’s name, an important identifying piece of personal information, can potentially be revealed from a digital photograph of the individual’s face by processing the image through a facial recognition system. A revealed name can then be connected to additional identifying and private pieces of information about the individual. Whether or not a facial recognition system successfully identifies an individual depends largely on the quality of the photo being processed. Quality factors such as resolution, luminosity, angle, and positioning each add a layer of difficulty to recognition that accumulates, and facial recognition systems must acquire the ability to trivialize these variables for efficient recognition performance. Even under ideal conditions, challenges arise with recognition due to the diversity of human faces and their expressions. Facial recognition can be implemented with a variety of different methods, each with their own strengths and weaknesses regarding complexity, detection capability, and recognition accuracy. Deep learning has been adapted in facial recognition technology in an attempt to improve the recognition accuracy of facial recognition systems over traditional non-learning methods. The use of a deep learning convolution neural network for a facial recognition system increases recognition accuracy over a traditional linear discriminant analysis approach which raises privacy concerns for people scanned with technology equipped with facial recognition capabilities. As technology progresses and facial recognition systems improve and outperform prior methods, the ethical implications surrounding privacy must be explored and addressed. An individual being scanned with a facial recognition system will not typically be aware of such scan being conducted if the scan is executed through public surveillance.

**Background**

Public surveillance systems for security purposes are extremely common. An individual conducting day-to-day interactions in public will typically be recorded by some sort of surveillance system and leave behind a passive digital footprint during their travels. However, public surveillance systems can have limits in ability if no processing of the video footage is conducted to provide meaningful information outside of human visual analysis. An individual usually gets identified in security footage only if the system operators recognize them. Limitations of surveillance systems could soon change with a rise in facial recognition technology. Robert N. Charette (2018), an international authority on information technology and systems risk management, states the cost of entry for facial recognition technology is dropping which attracts law enforcement agencies of varying size to the technology (para.10). A drop in cost leads to a rise in accessibility, and smaller agencies with less funding could potentially adapt facial recognition technology to process security footage with. Surveillance systems with facial recognition capabilities could become more widespread and commonplace, even in small, rural areas. Widespread use of facial recognition systems paired with an increase in performance and accuracy of recognition pose concerns of privacy. To improve facial recognition systems, Sulis Setiowati, Zulfanahri, Eka Legya Franita, and Igi Ardiyanto (2017), researchers in the field of artificial intelligence and machine learning from Gadjah Mada University, states in a conference paper that using a convolution neural network (CNN) for deep learning approaches yield the best results in regards to accuracy over traditional non-deep learning methods (p.1). The conference paper depicts deep-learning as, “very efficient to use in predicting . . . known or unknown data” (p.4). Not only can deep learning identify individuals from images containing all necessary data for recognition, but obstructed face images could still reveal identification with the use of deep learning. CNNs are powerful tools in the field of deep learning, and when applied to facial recognition, offer impressive results. However, non-deep learning methods for facial recognition are still efficient and reliable such as the linear discriminant analysis method.

**Precedents and Related Work**

Linear discriminant analysis (LDA) is a linear transformation technique in the field of linear algebra for dimension reduction of a matrix. Since digital images can be represented as a matrix of pixels, each pixel representing a specific color value, linear algebra computations can be applied to the matrix of a digital face image for facial recognition purposes. Raywut Ketsuwan and Praisan Padungweang (2017), researchers from King Mongkut's University of Technology Thonburi, states that LDA is used for facial feature extraction (p.1). Ketsuwan and Padungweang explains the process as transforming the input data into a new vector space that has a high discrimination among classifications (p.1). Ketsuwan and Padungweang continue by differentiating between within-class scatters that are minimalized and between-class scatters that are maximized in the newly transformed space (p.1). Figure 1 displays what the points of a scatter plot output look like using LDA on three different objects (Raschka, 2014). Between-class scatters are the three distinct scatter plots separated by color and shape that distinguish each individual object. Within-class scatters are scatter plots consisting of points of the same color that all relate to the specific object being identified. The goal behind LDA is to maximize the separation between different objects’ scatter plots so distinct classifications can be observed, while simultaneously minimalizing the separation distance among the points of an object’s scatter plot.

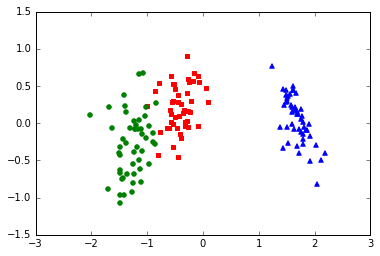


Figure 1. Scatter plot produced by linear discriminant analysis of three objects

The resulting projections from linear discriminant analysis computations can be applied to facial recognition to classify and identify faces in a digital image. An individual’s face will produce a unique scatter plot when the LDA algorithm is applied to a digital image of their face. LDA is a holistic approach which implies the overall image is evaluated for classification instead of individual discrete sections. LDA still offers efficiency as a simple method and remains relevant for facial recognition software. Convolution neural networks for deep learning pursue an oppossing approach when compared to the LDA method in an attempt to improve recognition accuracy.

**Support**

**Neural Networks**

Neural networks consist of interconnected nodes grouped in parallel layers. The input and output layer are each separated by a series of hidden layers. Data fed through the input layer traverses a specific path among the nodes of the hidden layers as a result of a weighted decision process. The decision process depends on the goal of the neural network, and the weights are uniquely tailored by training the neural network with training data sets. The final hidden layer connects to an output layer to retrieve a specific result. Figure 2 illustrates a graphical representation of a basic neural network (Jewmaidang & Tunchanok, 2018). The number of hidden layers, nodes within each layer, and edges between nodes will vary among different structured network architectures.

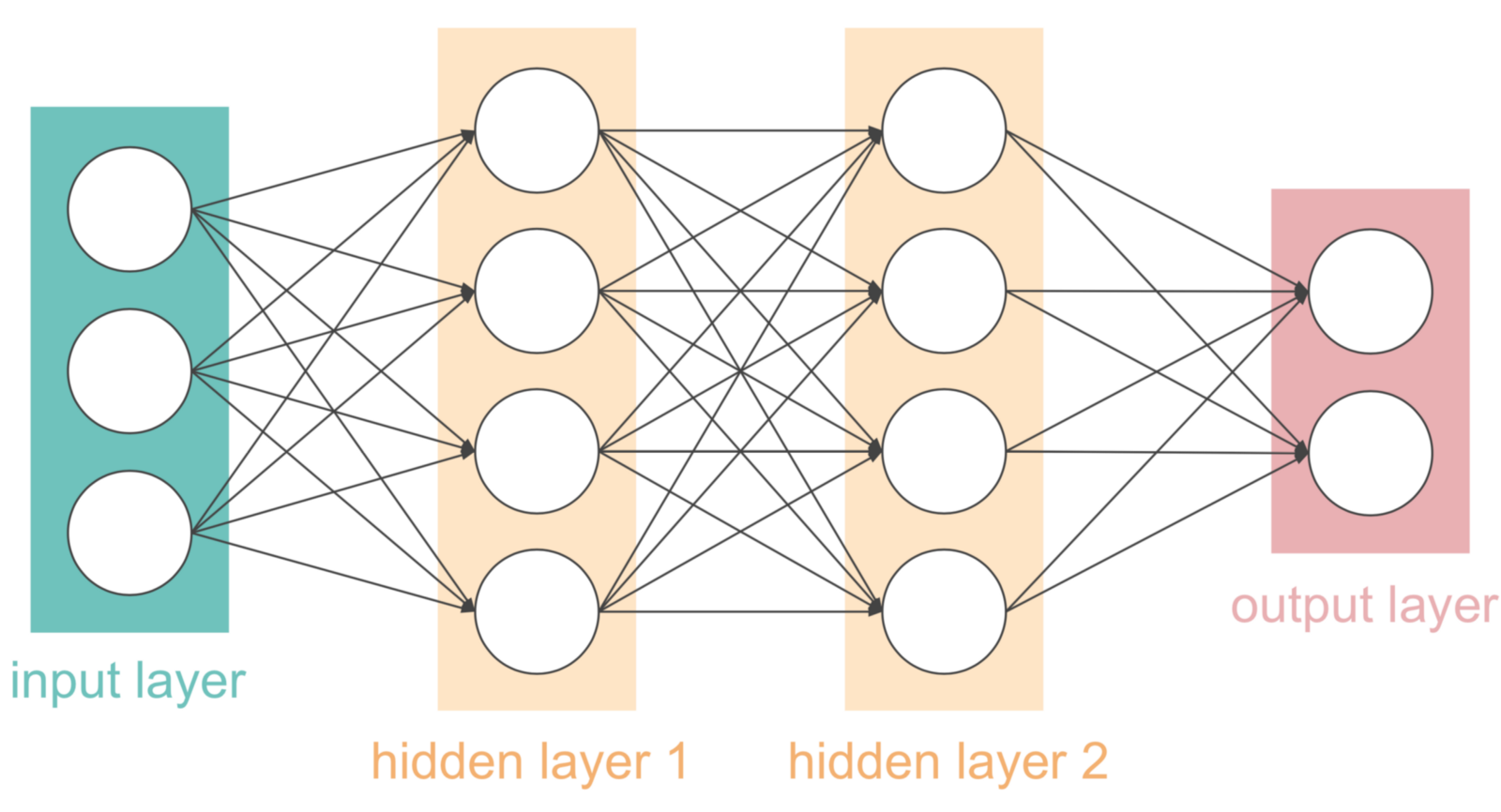


Figure 2. Architecture of a simple neural network

Convolution neural networks are a type of neural network with the addition of convolution layers. For deep learning, the number of hidden layers and nodes in a CNN are larger for higher complexity problems like facial recognition.

**Convolution Neural Network Architecture**

Setiowati, Zulfanahri, Legya Franita, and Ardiyanto’s conference paper states CNNs use the convolution layer for a process called convolution. The paper describes this process as two functions in which the output of one function is fed through the other function repeatedly to extract features from an input image in the form of linear transformation data (p.4). An image feature is any independent, recognizable pattern in an image that can be used for recognition purposes. The convolution process is beneficial for facial recognition because an image of a face has multiple components to be processed for recognition, such as a person’s mouth or their eyes. Instead of computing the entire image for facial recognition, like the LDA approach, individual features of the image are extracted for computing in a CNN. The process of convoluting an image is done to simplify the data for recognition. Setiowati, Zulfanahri, Legya Franita, and Ardiyanto explains the next layer, the subsampling layer, which uses a technique called max pooling to reduce the size of the image in which output features extracted from the convolution layer are divided into multiple small grids; the highest value of each small grid is then organized into a new matrix. Setiowati, Zulfanahri, Legya Franita, and Ardiyanto describe the next layer, the full connection layer, which is a layer responsible for performing a transformation on the data dimension from the subsampling layer so the data can be linearly classified. The conference paper concludes the description of CNN architecture with the output layer (p.4). A CNN has intricate architecture with various layers that all work together to process an image. When applied to facial recognition, images of human faces can be fed through a CNN and offer efficient results.

**The Power of Convolution Neural Networks**

Convolution neural networks are becoming a common choice for facial recognition systems for their efficiency. Setiowati, Zulfanahri, Legya Franita, and Ardiyanto’s conference paper outlines results regarding research focused around comparing different facial recognition methods using two publicly available image sets, ORL and YALE. The research results concluded the LDA method achieved 87.7% and 85.2% rate of recognition on the ORL and YALE image sets respectively, and the deep learning CNN method achieved 92.6% and 93.3% rate of recognition (p.5). The CNN outperformed the LDA in both image sets, achieving a 4.9% improved rate of recognition on the ORL image set and an 8.1% improved rate of recognition on the YALE image set. Although the rate of recognition is largely influenced by the image set used, the trend among the ORL and YALE image sets are in favor of deep learning CNNs offering improved efficiency over the non-deep learning LDA method. A deep learning CNN proves itself as a desirable method for its power and efficiency.

Deep learning CNN architectures must go through a training process to dial in potential. The more unique data the network has to train with, the higher potential the network will have to learn and improve recognition accuracy. Hu, Yang, Yi, Kittler, Christmas, Li, and Hospedales (2015), Hu, an honorary assistant professor at Queen’s University Belfast, and Yang, a research assistant at University of Edinburgh, state in a conference paper that Facebook has a team of AI researchers who designed and trained an 8 layer CNN called DeepFace. The paper outlines that researchers for DeepFace had access to 4,000 subjects and 120 million images to train the network with (pp.385-386). A deep learning CNN system like DeepFace having access to such large training sets to increase performance is worrisome for privacy. DeepFace has gone through rigorous training, and the only factor barring other CNN systems from the same level of training is the availability of training data sets, but massive data sets could soon become more commonplace.

**Applied Facial Recognition**

With CNN’s popularity and efficiency on the rise for facial recognition, more systems may begin to adapt CNN architecture in the future with obtainable training data. Alessandro Acquisti, Professor of Information Technology and Public Policy at Carnegie Mellon University, states in his TED talk (2013) that not only are facial recognition systems improving, but the availability of facial image data is increasing as well because of how many pictures people share through the Internet on public domains (1:03). Deep learning CNNs are contributing to the improvement of facial recognition systems, and when paired with an increase in facial image data availability, pose a concern for privacy since larger training sets can increase CNN performance. Acquisti conducted an experiment on Carnegie Mellon University campus where students were asked to have their picture taken before filling out a survey. Acquisti used a facial recognition system connected to a database with hundreds of thousands of images scraped from Facebook profiles to cross reference the student’s photo with. Acquisti found that one in three students were identified by the system (2:35). The experiment successfully took an image of a person’s face to produce their name in the short amount of time it took the student to fill out a survey. The experiment was conducted using a reference database containing publicly available profile pictures from Facebook. The names produced from the experiment are also public pieces of data that are displayed next to an individual’s Facebook profile picture. Privacy risks surrounding facial recognition technology extend beyond information individuals publicize. Personal private information could potentially be revealed from an individual’s name captured by facial recognition software when in the hands of law enforcement.

Facial recognition technology can be utilized with nothing more than an Internet connection. Acquisti comments on the increase in accessibility of cloud computing services (1:03). Cloud computing provides use of remote servers accessible through the Internet to process data. Cloud computing links a user to computing power they may be unable to achieve with their own personal computer. Facial recognition systems are already being commercialized and hosted on cloud based services as seen in Charette’s article with Amazon’s deep learning Rekognition software service (para.1). Law enforcement agencies of all sizes have the ability to process surveillance data for facial recognition without housing the hardware needed to do so. When surveillance footage and facial recognition reveal a name, law enforcement agencies have access to a wide variety of databases to search for that name and reveal additional information about the identified individual. Individuals have no way of knowing if they are being scanned with facial recognition technology, but a conclusion could be drawn that law enforcement agencies only have interest in powerful facial recognition systems for identifying criminals, and ordinary citizens should withdraw concern.

**Regulation of Facial Recognition Software**

Privacy invasive systems, like facial recognition software, require regulation to ensure public safety. In his article, Charette points out the exploitation of facial recognition technology by government agencies, “In 2016, the U.S. Government Accountability Office (GAO) published a report declaring that the FBI didn’t always adhere to federal privacy laws and policies affecting [automatic facial recognition] use” (para.5). Even with regulations in place, facial recognition technology can attract abusive behavior from government agencies. With CNNs on the rise for facial recognition to improve accuracy, the privacy of people scanned with facial recognition technology is at risk if agencies are not addressing the ethical implications with the technology being used. Charette states in his article that facial recognition researchers claim facial recognition software has the potential to determine a person’s sexual orientation (para.11). A deep learning CNN can be trained with facial images of people who align with a certain sexual orientation in order to use the system to identify sexual orientation, a private piece of information an individual may or may not be willing to share, of anyone scanned with the technology. Trainable deep learning CNNs harness the power to go beyond basic recognition that cannot be achieved through non-deep learning methods like LDA.

**Conclusion**

People in public interact with surveillance systems on a daily basis, leaving behind a passive digital footprint which individuals may or may not be aware of. With systems like deep learning CNNs improving facial recognition accuracy over non-deep learning methods, an individual’s face could potentially be scanned with facial recognition technology without their knowledge. Law enforcement agencies of all sizes having access to facial recognition software allows processing of surveillance footage with relative ease due to cloud computing. Even though law enforcement will typically focus their interest on a criminal, innocent bystanders may have their faces scanned simply to verify that they are not the criminal being searched for. Data produced from facial recognition scans must be handled with sensitivity, but government agencies like the FBI have been exposed for their lack of care over sensitive data in regards to facial recognition, even with regulations in place meant to prevent the exploitation of data from occurring.

CNNs will continue to evolve and improve beyond capabilities seen in recent research. CNNs may eventually achieve near human-like performance and raise the bar for facial recognition software. A single photo of an individual may be abstracted down to nothing but a matrix of pixels, but to a facial recognition system, those pixels can reveal more information than just the look on an individual’s face. A digital photograph of an individual can potentially link to their name, which then links to digital records and information about them. If a CNN is trained to detect a certain quality in an individual through their facial expressions, facial recognition systems may begin to make characteristic predictions, like sexual orientation, to infringe on an individual’s privacy without their knowledge. What if a deep learning CNN is created and trained with facial images from criminal databases so law enforcement could potentially predict whether a person is a criminal? Unethical profiling could occur on individuals before they even break the law. As deep learning CNNs aid in the evolving world of facial recognition, the ethical guidelines of facial recognition systems must evolve alongside the technology itself to maintain the integrity of public safety and privacy.

References

Acquisti, A. (2013, June). What will a future without secrets look like? [Video file]. TED. Retrieved from https://www.ted.com/talks/alessandro\_acquisti\_why\_privacy\_matters

Charette, R. N. (2018, May 29). Automated facial recognition: Menace, farce, or both? *IEEE Spectrum*. Retrieved from https://spectrum.ieee.org/riskfactor/computing/it/automated-facial-recognition-menace-farce-or-both

Hu, G., Yang, Y., Yi, D., Kittler, J., Christmas, W., Li, S. Z., Hospedales, T. (2015). *When face recognition meets with deep learning: An evaluation of convolutional neural networks for face recognition*. Paper presented at the 2015 IEEE International Conference on Computer Vision Workshop, Santiago, Chile. doi:10.1109/ICCVW.2015.58

Ketsuwan, R., Padungweang, P. (2017, July). *A linear discriminant analysis using weighted local structure information*. Paper presented at the 2017 14th International Joint Conference on Computer Science and Software Engineering, Nakhon Si Thammarat, Thailand. doi:10.1109/JCSSE.2017.8025907

Jewmaidang, K., & Tunchanok, P. (2018). *Interns explain basic neural network*. Retrieved from https://cdn-images-1.medium.com/max/1600/1\*GM2i5bVgcmlQOEN4lP81JQ.png

Raschka, S. (2014). *Linear discriminant analysis.* Retrieved from https://sebastianraschka.com/images/blog/2014/linear-discriminant-analysis/2014-08-03-linear\_discriminant\_analysis\_154\_1.png

Setiowati, S., Zulfanahri, Legya, E. F., Ardiyanto, I. (2017, October). *A review of optimization method in face recognition: Comparison deep learning and non-deep learning methods*. Paper presented at the 2017 9th International Conference on Information Technology and Electrical Engineering, Phuket, Thailand. doi:10.1109/ICITEED.2017.8250484