

# Research on facial recognition system based on deep learning

**Yiming Yao**

Khoury College of Computer Science, Northeastern University, Boston, America

yao.yim@northeastern.edu

**Abstract.** This essay explores the integration of deep learning techniques in facial recognition, delving into the advancements, challenges, and prospects in future of this transformative technology. Deep learning, a subset of artificial intelligence, has revolutionized facial recognition by enabling automatic feature learning and robust representations from raw facial data. This paper examines the key components of deep learning in facial recognition, including feature extraction, facial alignment, and face recognition. It highlights the advantages of using deep learning over traditional methods by analysing related research. It also addresses the challenges deep learning faces in facial recognition, such as data bias, privacy concerns, and adversarial attacks. The essay also explores the promising future of deep learning in facial recognition, with ongoing research aiming to improve robustness, fairness, and privacy preservation.

**Keywords:** facial recognition, deep learning, feature extraction.

## 1. Introduction

A face recognition system is a technological system that compares the face information extracted from an image or video with the information in the database to see if the results match. Due to its practicality and convenience, facial recognition technology has been widely used worldwide, including unlocking, security monitoring, etc. The following steps can achieve real-time facial recognition [1]. First, the face will be separated from the image, which is face detection. Then, the image of the face will be aligned by change the properties of the image like size and grayscale. The third step is extracting the facial identities from the image, then compare them with the counterparts from database, which is the last step [1]. And deep learning is a subset of machine learning, a branch of artificial intelligence that has shown remarkable capabilities in various complex tasks, including facial recognition. At its core, deep learning involves training artificial neural networks with multiple layers, also known as deep neural networks, to learn hierarchical representations of data [2]. The term "deep" refers to the depth of the network, which can have tens or hundreds of layers, enabling it to capture intricate patterns and relationships within the input data. In this subsequent essay, the authors will explain the role of each step of face recognition and how deep learning can help in this step and how it is an improvement over previous approach. Also, this side of the paper will talk about the problems of deep learning in face recognition as a research area and the future prospects of deep learning.

## 2. Facial recognition

### 2.1. Deep learning

Deep learning is a major innovation for face recognition systems because of the flexibility, adaptability, and highly accuracy. Also, deep learning has helped in all aspects of face recognition systems. Facial detection, a computer vision method, recognizes human faces' presence and positions in images or video frames. It serves as a fundamental component in facial recognition systems and other applications that heavily rely on face detection, therefore, the accuracy of the detection is critical [3]. Before deep learning is widely used in facial detection, people used some methods like Haar cascades and Viola-Jones algorithm. However, the limitations exist, including difficulty in adapting to complex scenarios, required manual intervention and so no [4]. With the ability to learn hierarchical features from raw data, deep learning models, especially convolutional neural networks (CNNs), have revolutionized facial detection [5]. These models can learn to capture complex patterns and variations in facial images, enabling them to handle the challenges posed by real-world scenarios [3].

### 2.2. Facial detection with deep learning

Data collection is the initial and essential step to training a useful model. A large dataset of labeled facial images is gathered, covering various individuals with diverse facial appearances, poses, lighting conditions, and expressions [1]. The dataset is crucial for training the deep learning model effectively. The gathered facial images undergo several preprocessing procedures, including resizing, normalization, and data augmentation [6]. Data augmentation methods introduce alterations to the images, such as rotations, flips, and brightness adjustments, to enhance dataset diversity and promote better generalization of the model. Convolutional Neural Networks (CNNs) are the most used architecture because they can effectively learn hierarchical features from images [5]. The crucial stage involves training the chosen deep learning model on the preprocessed dataset. Throughout the training process, the model gains the ability to extract significant facial features from images and associate them with their respective identities. This training process entails forward and backward propagation to optimize the model's parameters (weights and biases) using a designated loss function [3]. After training, the model undergoes evaluation using a distinct validation dataset to gauge its performance, and if needed, it is fine-tuned. Fine-tuning may entail adjusting hyperparameters or retraining the model on a smaller dataset to enhance its performance in handling specific tasks [7].

### 2.3. Facial alignment with deep learning

Facial alignment is a crucial step in facial recognition system as it ensures that the detected faces are correctly oriented and aligned for accurate feature extraction and matching [1]. Traditional methods for facial alignment relied on handcrafted feature points or landmarks, but they often struggled with variations in pose, expressions, and occlusions [3]. However, with deep learning, facial alignment has significantly improved precision and efficiency. Facial alignment encounters difficulties due to variations in pose, where the face's angle and orientation can vary significantly. Moreover, facial expressions and outside objects, like glasses, facial hair, or hands covering parts of the face, masks, present challenges in accurately aligning the facial features [3]. Conventional approaches often need extensive manual adjustments and may struggle to address these issues effectively, resulting in low-optimal alignment outcomes. Deep learning methods, particularly convolutional neural networks (CNNs), have succeeded in facial alignment. Using extensive datasets and hierarchical feature learning, CNNs can autonomously acquire representations that grasp the intrinsic diversities present in facial images, rendering them suitable for complex alignment tasks [8].

### 2.4. Facial extraction with deep learning

While feature extraction is also important in facial recognition system because it can represent the feature of faces in a compact and discriminative way [1]. Deep learning has revolutionized feature extraction, enabling extracting high-level and semantically meaningful features from raw facial data

[8]. Facial recognition systems capture unique and informative facial characteristics to accurately distinguish one individual from another. Traditional methods used handcrafted features, such as Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP). While these approaches showed some success, they could not learn complex and abstract representations that are essential for robust facial recognition across diverse conditions. While by using deep learning model, like convolutional neural networks (CNNs), has shown to be more effective and accurate in feature extraction. Convolutional neural networks (CNNs) are composed of multiple layers, with each layer progressively learning more abstract representations of the input data [7].

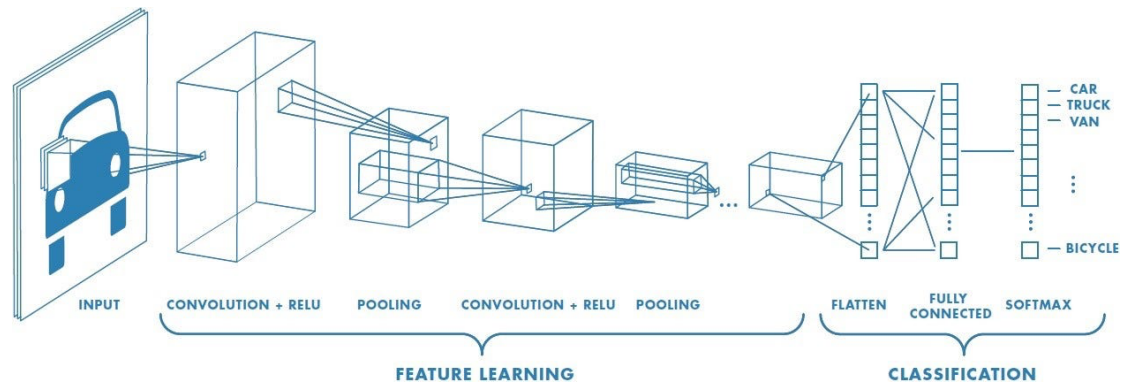
Regarding facial recognition, CNNs can hierarchically detect and represent various facial characteristics, like eyes, nose, and mouth. One of the key advantages of deep learning is the availability of pre-trained models. Researchers have trained large-scale CNNs on massive datasets, such as ImageNet, which contain millions of images from various categories. These pre-trained models have learned to extract general-purpose features, and their learned representations can be transferred and fine-tuned for specific tasks, including facial recognition [8].

### *2.5. Facial recognition with deep learning*

Various similarity metrics can be employed to assess the resemblance between the feature vector of the detected face and the feature vectors stored in the database. Commonly used metrics encompass Euclidean distance, cosine similarity, and the L1 or L2 norm, which quantify the likeness or dissimilarity between two feature vectors [9]. Subsequently, the system computes the similarity score between the detected face's feature vector and each entry in the database. This score is then compared against a predetermined threshold to ascertain whether the detected face corresponds to any known identity. The threshold determines the required similarity level for a positive match. If the similarity score surpasses the threshold, the system confirms it as a match; otherwise, the face is considered as unidentified or unrecognized [6]. Based on the similarity scores and thresholding, the facial recognition system decides the identity of the detected face. In the event of a match in the database, the system identifies the individual and provides the relevant outcome. If no match is found, the system may categorize the face as an unknown identity or request further verification [6].

### *2.6. One research about CNN*

In one area, Feedforward Neural Networks (FFNN) and Convolutional Neural Networks (CNN) are used to classify facial features in a wide range of categories, with a particular focus on identifying one person from more than 5,000 categories [7]. For FFNN, the goal output is Boolean, whereas CNN uses the real data set to learn and recognize it. The study included the collection of real data from 1, 25 pictures of a test subject showing various expressions and backgrounds. Additionally, 13,530 images from various subjects were sourced from the "labelled faces in the wild" dataset. A random mixture of pictures of the experimental subjects was used to produce over 14,000 pictures. The main goal was to successfully recognize the face of the target in this large dataset. In this paper, a variety of facial features were explored, and a local normalization histogram of orientation gradient (HOG) was demonstrated in comparison with others such as wavelets. The Viola-Jones method is used to detect facial features. Convolution is defined as a method for generating characteristic maps by means of a filter or a kernel. Each convolution employed different filters, generating distinct feature maps. These feature maps were aggregated to form the layer's output, which underwent activation functions and subsequent pooling layers to reduce dimension.



**Figure 1.** Process of Convolutional Neural Network [10].

The last step involved converting the output into a one-dimensional feature vector for classification (Fig. 1). The VGG-16 model consists of 16 convolutional layers, 5 max pooling and 3 fully-coupled layers, and finally a Soft Map layer. The ReLU is activated for the hidden layer, and the full connection layer is subjected to dropout normalization. The precision of CNN was 99.16%, and that of FFNN was 92.80% in experiment assessment [7].

### 3. Analysis of deep learning

#### 3.1. Defects and challenges

The progress of deep learning in facial recognition has brought about a transformative impact on the field, although it is not without its difficulties [11]. A key issue is data bias, where biased training data can lead to inaccurate and inequitable outcomes, particularly for underrepresented groups. Additionally, privacy concerns arise due to the sensitive biometric information involved in facial recognition, and the potential misuse or unauthorized access to this data gives rise to ethical and legal dilemmas. Adversarial attacks present a security risk, as even imperceptible changes to facial images can cause misclassification. Moreover, deep learning models may encounter challenges in accurately performing under real-world conditions with varying lighting, poses, and occlusions, thus affecting their robustness. The resource-intensive training requirements and ethical concerns surrounding surveillance and discrimination further complicate the implementation of facial recognition technology [12].

#### 3.2. Future prospects

Nonetheless, despite these challenges, the future of deep learning in facial recognition holds promise. Ongoing research endeavors focus on enhancing model robustness and fairness, addressing biases, and advancing techniques that preserve privacy [12]. With hardware acceleration and the integration of multimodal recognition, real-time and dependable deployment in diverse environments becomes feasible. Emphasizing interdisciplinary collaboration among computer scientists, ethicists, legal experts, and policymakers will be crucial in shaping a future where facial recognition can be responsibly harnessed to bring about positive societal impact while upholding individual rights and privacy [4].

### 4. Conclusion

In conclusion, deep learning has made huge advancement for facial recognition system, revolutionizing the field with its automatic feature learning, improved accuracy, and robust representations. By enabling the extraction of complex facial attributes from raw data, deep learning models have overcome the limitations of traditional methods, excelling in real-world scenarios with varying poses, lighting, and occlusions. The advantages of deep learning in facial recognition extend

to scalability, adaptability, and real-time performance, making it a preferred choice for applications ranging from access control to surveillance.

However, alongside these remarkable advancements, the essay has also addressed the challenges faced by deep learning faces in facial recognition, such as data bias, privacy concerns, and potential adversarial attacks. Ethical considerations must be at the forefront of development and deployment, ensuring that facial recognition technology is used responsibly and with respect for individual rights and privacy.

The future of deep learning in facial recognition holds great promise, as ongoing research endeavors strive to enhance model robustness, mitigate biases, and preserve privacy. Multimodal recognition integration and interdisciplinary collaboration will play pivotal roles in shaping this technology's trajectory, as computer scientists, ethicists, legal experts, and policymakers collaborate to ensure ethical and equitable implementation.

As we move forward, it is imperative to strike a delicate balance, harnessing the potential of deep learning in facial recognition while addressing societal concerns and ethical implications. By doing so, we can build a future where facial recognition technology contributes positively to various domains, empowering us with sophisticated tools that respect privacy, foster fairness, and promote responsible use. Through continuous advancements and responsible deployment, deep learning and facial recognition will undoubtedly play a transformative role in shaping our digital society for years to come.

## References

- [1] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, *Sensors*, vol. 20, no. 2, p. 342, 2020.
- [2] X. Hao, G. Zhang, and S. Ma, vol. 10, no. 03, pp. 417–439, 2016.
- [3] K. H. Teoh, R. C. Ismail, S. Z. M. Naziri, R. Hussin, M. N. M. Isa, and M. Basir, in *Journal of Physics: Conference Series*, IOP Publishing, 2021, p. 012006.
- [4] Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, *Electronics*, vol. 9, no. 8, p. 1188, 2020.
- [5] X. Qi, C. Wu, Y. Shi, H. Qi, K. Duan, and X. Wang, *Computational Intelligence and Neuroscience*, vol. 2023, 2023.
- [6] H. Wang and L. Guo, *IEEE*, 2021, pp. 540–546.
- [7] N. S. Singh, S. Hariharan, and M. Gupta, in *Advances in Data Sciences, Security and Applications: Proceedings of ICDSSA 2019*, Springer, 2020, pp. 375–382.
- [8] E. M. Onyema, P. K. Shukla, S. Dalal, M. N. Mathur, M. Zakariah, and B. Tiwari, *Journal of Healthcare Engineering*, vol. 2021, 2021.
- [9] [C. Cruz, L. E. Sucar, and E. F. Morales, in *2008 8th IEEE international conference on automatic face & gesture recognition*, IEEE, 2008, pp. 1–6.
- [10] S. Saha, Dec. 15, 2018. <https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/> (accessed Aug. 11, 2023).
- [11] P. Kaur, K. Krishan, S. K. Sharma, and T. Kanchan, *Medicine, Science and the Law*, vol. 60, no. 2, pp. 131–139, 2020.
- [12] S. Singh and S. Prasad, *Procedia computer science*, vol. 143, pp. 536–543, 2018.