# Face Recognition: From Traditional to Deep Learning Methods

Siyu Huang
The university of Adelaide
a1810323

## Abstract

*Face recognition technology is a biometric technology, which is based on the identification of facial features of a person. People collect the face images, and the recognition equipment automatically processes the images. The paper introduces the application of face recognition, as well as the traditional and deep learning-based approach for face recognition.*

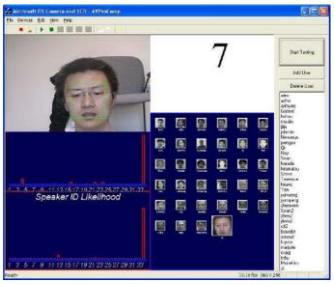
## 1. Introduction

Face recognition refers to the technology capable of identifying or verifying the identity of subjects in images or videos and is performed to extract the main features of the face from the input image, and then the detected features are compared to the data set image. (Facial recognition system - Wikipedia, 2022) The reason face recognition is more appealing and wider used than other biometric modalities is due to its non-intrusive nature. For example, fingerprint recognition requires users to place a finger in a sensor, iris recognition requires users to get significantly close to a camera. In contrast, modern face recognition only requires the users to be within the field of view of a camera. Hence, tremendous potential can be revealed from face recognition as it can work in an environment where the users are not aware or expected to cooperate. Face recognition has a large number of applications, including security, person verification. In this paper, a traditional approach and a deep learning-based approach will be proposed and their advantages and limitations will be discussed respectively.

### 1.1. Face recognition applications

*1.1.1 Airport security system*

Today more than ever, security is a primary concern at airports and for airline personnel and passengers. Airport security systems that use face recognition technology have been implemented at many airports around the globe. Figure 16.3 diagrams a typical airport security system that Fig. 16.2.




employs face recognition technology.

*1.1.2 Antique photo verification*

It is of great value for historians, biographers, and antique collectors to verify whether an antique photo of a person is genuine, given a true photo taken when that person is much older. The age difference and sometimes the low quality of the antique photo poses a great challenge for the face recognition systems.

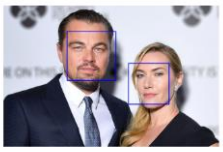
### 1.2. Approaches for face recognition

*1.2.1 Traditional approach*

The proposed traditional face recognition system composes of the following building blocks:


1) Face detection.

Once the facial images have been captured by the camera, face detection can be accomplished by comparing the physical appearance of the user in the image processing. (Facial recognition system - Wikipedia, 2022) The face detector will find the position of the faces in an image and return the coordinates of a bounding box for each one of them. This is illustrated in Figure a.



(a)

2) Face alignment.

After the face detection, face alignment will scale and crop face images, in the same way, using a set of reference points located at fixed locations in the image. This process requires finding a set of facial landmarks using a landmark detector and, in the case of a simple 2D alignment, finding the best affine transformation that fits the reference points. Figures b and c show two face images aligned using the same set of reference points.

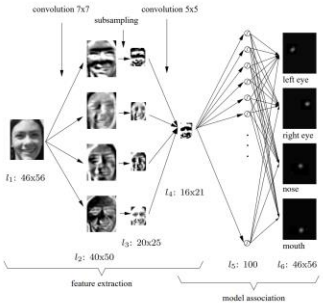
(b) (c)

3) Face representation.
At the face representation stage, the pixel values of a face image are transformed into a compact and discriminative feature vector, also known as a template. Ideally, the faces with the same subject can be mapped to similar feature vectors.

4) Face matching.
Lastly, in the face matching building block, two templates will be compared in order to produce a similarity score which indicates the likelihood that they belong to the same subject.

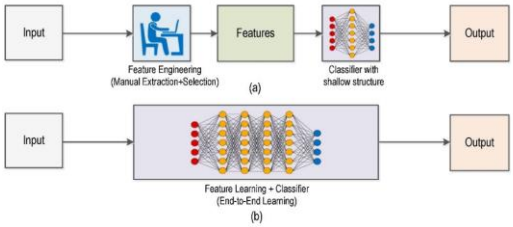
### *1.2.2 Deep learning-based approach*

The proposed face recognition based on convolutional neural networks needs to collect a large number of pictures for the computer to learn first. CNN designed in this paper contains the following layers of structure which consist of six layers, where the first layer is the input layer, the three following layers are convolutional and sub-sampling layers, and the last two layers are standard feedforward neuron layers. The aim of the system is to learn a nonlinear function that transforms a raw input face image into desired output feature maps where facial features are highlighted.




## 1.3. Comparison between two approaches

Face recognition techniques have shifted significantly over the years. Traditional methods relied on hand-crafted features, such as edges and texture descriptors, combined with machine learning techniques. Recently, traditional face recognition methods have been superseded by deep learning methods based on convolutional neural networks (CNNs). The main advantage of deep learning methods is that they can be trained with very large datasets to learn the best features to represent the data. Because of this, the workflow of the CV engineer has changed dramatically where the knowledge and expertise in extracting hand-crafted features have been replaced by knowledge and expertise in iterating through deep learning architectures as depicted in the below figure.




### *1.3.1 Limitations of traditional approach*

The traditional approach to face recognition involves using well-established CV techniques as feature descriptors for object detection, for example, SIFT, SURF, BRIEF, etc. A step in the traditional approach called feature extraction was carried out for image classification before the emergence of DL. Several CV algorithms, such as edge detection, corner detection, or threshold segmentation may be involved in this step. One challenge with this traditional approach is to determine which features in a given image are important, and as the number of classes to classify increases, feature extraction becomes harder. The traditional approach mostly relies on the CV engineer's judgment to decide which features can best describe different classes of objects, which may cause undesired performance and mistakes.

### *1.3.2 Advantages of traditional approach*

Traditional CV techniques can sometimes solve a problem much more efficiently than DL. Algorithms like SIFT and even simple color thresholding and pixel counting algorithms are not class-specific, that is, they are very general and perform the same for any image. In contrast, features learned from a deep neural net are specific to the training dataset which, if not well constructed, probably won't perform well for images different from the training set.

### *1.3.3 Advantages of Deep Learning approach*

DL introduced the concept of end-to-end learning where the machine is just given a dataset of images that have been annotated with what classes of the object are present in each image. Thereby a DL model is 'trained' on the given data, where neural networks discover the underlying patterns in classes of images and automatically work out the most descriptive and salient features with respect to each specific class of object for each object. Hence, DL approaches may achieve substantially better accuracy than traditional approaches.

### *1.3.4 Limitation of Deep Learning approach*

Since a DL model needs to be trained on a large dataset, and multiple layers consist of CNN to discover the underlying patterns in image classes, billions of additional math operations are used and the requirement for processing power is heavily increased.

## 1.4. CONCLUSIONS

Two approaches have been proposed in this paper for face recognition systems. Although the traditional approach based on hand-engineered features can sometimes be more efficient on non-class-specific problems, Deep learning-based approaches have become the standard for face recognition due to the significant accuracy improvement achieved over other types of methods. One thing that needs to be careful with the DL approach is collecting large amounts of labeled face images is expensive, and very deep CNN architectures are slow to train and deploy.

Word Count: 867

## References


[1] M. Lyons, A. Plante, S. Jehan, S. Inoue, and S. Akamatsu. Avatar creation using automatic face recognition. In Proceedings, ACM Multimedia 98, pages 427–434, 1998.

[2] En.wikipedia.org. 2022. Facial recognition system - Wikipedia. [online] Available at: <https://en.wikipedia.org/wiki/Facial_recognition_system> [Accessed 5 June 2022].

[3] Y. Chan, S.-H. Lin, and S. Kung. Video indexing and retrieval. In Multimedia Technology for Applications. B.J. Sheu, M. Ismail, M.Y. Wang, and R.H. Tsai, editors. Wiley, New York, 1998.

[4] FaceGate. http://www.premierelect.co.uk/faceaccess.html.

[5] P. Hong, Z. Wen, and T. Huang. IFace: a 3D synthetic talking face. International Journal of Image

[6] and Graphics, 1(1):19–26, 2001.