Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 17 No. 23 (2023) | OPEN ACCESS

https://doi.org/10.3991/ijim.v17i23.40867

PAPER

# A Review of Deep Convolutional Neural Networks in Mobile Face Recognition

Jing Chi[1], Chin Kim On[1](✉), Haopeng Zhang[2], Soo See Chai[3]

[1]Faculty of Computing and Informatics, University Malaysia Sabah, Sabah, Malaysia

[2]School Information and Electrical Engineering, Hebei University of Engineering, Hebei, China

[3]Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, Sarawak, Malaysia

kimonchin@ums.edu.my

**ABSTRACT**

With the emergence of deep learning, Convolutional Neural Network (CNN) models have been proposed to advance the progress of various applications, including face recognition, object detection, pattern recognition, and number plate recognition. The utilization of CNNs in these areas has considerably improved security and surveillance capabilities by providing automated recognition solutions, such as traffic surveillance, access control devices, biometric security systems, and attendance systems. However, there is still room for improvement in this field. This paper discusses several classic CNN models, such as LeNet-5, AlexNet, VGGNet, GoogLeNet, and ResNet, as well as lightweight models for mobile-based applications, such as MobileNet, ShuffleNet, and EfficientNet. Additionally, deep CNN-based face recognition models, such as DeepFace, DeepID, FaceNet, and SphereFace, are explored, along with their architectural characteristics, advantages, disadvantages, and recognition accuracy. The results indicate that many scholars are researching lightweight face recognition, but applying it to mobile devices is impractical due to high computational costs. Furthermore, noise label learning is not robust in actual scenarios, and unlabeled face learning is expensive in manual labeling. Finally, this paper concludes with a discussion of the current problems faced by face recognition technology and its potential future directions for development.

**KEYWORDS**

computer vision, deep learning, deep convolutional neural network, face recognition, lightweight deep CNN models

## 1 INTRODUCTION

Face recognition is a biometric identification technology that utilizes physiological features to extract human facial features via a computer and authenticate identity based on these features [1]. With the continuous development of deep learning technology, especially Deep Convolutional Neural Network (DCNN), the field of face recognition has seen significant progress in recent years. As a non-contact and easy-to-implement technology, face recognition is not only applicable to mobile

Chi, J., Kim On, C., Zhang, H., Chai, S.S. (2023). A Review of Deep Convolutional Neural Networks in Mobile Face Recognition. *International Journal of Interactive Mobile Technologies (iJIM)*, 17(23), pp. 4–19. https://doi.org/10.3991/ijim.v17i23.40867

Article submitted 2023-04-26. Revision uploaded 2023-07-24. Final acceptance 2023-10-21.

devices, but it can also effectively mitigate the risk of contact transmission in the current context of pandemic prevention and control. This technology plays a critical role in enhancing security and surveillance capabilities by providing automated recognition solutions such as traffic surveillance, access control devices, biometric security systems, and attendance systems [2].

Face recognition involves three general steps: face detection, feature extraction, and face classification [3]. Face detection extracts information about facial position, size, and key point position from face images. Feature extraction uses a Convolutional Neural Network (CNN) to extract facial features. Face classification performs discriminative classification by comparing the extracted facial features with the feature information in the database [4]. Although the use of CNN has been proven effective in face recognition, there is still room for improvement. The performance of these models depends on the variety of algorithms used. This article focuses on describing and comparing commonly used CNN algorithms in face recognition, including their architectural characteristics, advantages, disadvantages, and recognition accuracy.

The remainder of this paper is organized as follows: In Section 2, we provide a brief overview of classic CNN models, including their architectural characteristics, advantages, and disadvantages. We also summarize their error rates in classification tasks. In Section 3, we focus on analyzing face recognition models and their architectural characteristics, advantages, disadvantages, and recognition rates in face datasets. Finally, we discuss the challenges faced by current face recognition technology and propose future development directions for face recognition.

## 2 DEEP CNN

In recent years, the development of deep learning has led to a transformation in the theoretical models used for face recognition. Traditional artificial feature extraction models have been replaced by CNN models with higher quality and accuracy [5] [6]. Face recognition is achieved by feeding a face image into the CNN model, which extracts multi-dimensional features that are then compared to the facial features stored in the database. The most commonly used DCNN models include LeNet-5 [7], AlexNet [8], VGGNet [9], GoogleNet [10], and ResNet [11]. In addition, there are also lightweight CNNs, such as MobileNet [12][13][14], ShuffleNet [15][16], and EfficientNet [17][18], which will be briefly explained in the following subsections.

### 2.1 LeNet-5

The LeNet-5 model, proposed in 1998, is one of the earliest CNN models, comprising of three convolutional layers, two pooling layers, and two fully connected layers, using the Sigmoid activation function [7]. Thanks to its innovative architecture, LeNet-5 played a crucial role in advancing deep learning. Its multiple layers of convolution and pooling operations paved the way for the development of more complex and powerful CNN models. Additionally, LeNet-5 was designed specifically for handwritten digit recognition, a challenging task at the time, and achieved high accuracy on this task. Its success demonstrated the potential of CNNs for solving complex pattern recognition problems and inspired further research in the field. However, technological limitations, particularly hardware processing speed and stability, hindered the widespread adoption of CNN models, making it difficult for researchers to extend and apply these models further [19].