Journal of Computer Science

Original Research Paper

Facial Recognition and Discovery Using Convolution Deep Learning Neural Network

Abdel Karim Mohamed Baareh

Department of Applied Science, Ajloun University College, Al-Balqa Applied University, Ajloun, Jordan

Article history
Received: 27-03-2024
Revised: 30-07-2024
Accepted: 03-08-2024

Email: baareh@bau.edu.jo

Abstract: Facial recognition is a critical and well-established topic that has caught the interest of researchers across various application areas. Face detection and identification are critical components of human detection systems. Researchers in different fields used numerous techniques to recognize human faces in various positions considering several features, many results were also obtained. In this research, we aim to use a well-known method called a Deep Learning lightweight Convolutional Neural Network (CNN) for image recognition and detection techniques to contribute to solving the facial recognition problem. We used 280 FEI face images to recognize images of twenty persons from fourteen unique classes. Training and testing samples were considered. The proposed CNN model consists of only 16 layers, making it more portable and reliable for real-world applications. CNN shows outstanding performance, completing a training accuracy result of 97.73% and a testing accuracy result of 83.33%.

Keywords: Facial Recognition, Convolutional Neural Network, Deep Learning

Introduction

Designing and creating systems for detecting and recognizing faces has been a popular research topic in computer vision for years (Bower and Karlin, 1974). The difficulty of feature selection in recognition models has spurred investigations into faster and more accurate methods. Facial recognition has several practical applications. In the 1960s, Woody Bledsoe and his colleagues used computers to detect human faces (Bledsoe, 1968, 1966a-b; 1964; Bledsoe and Chan, 1965). Their primary work," man-machine," had humans recognizing facial features before computers could analyze faces (Nils, 2009).

The Army Research Laboratory and Defense Advanced Research Project Agency initiated a program in Facial Recognition to support this vital domain. The technologies created under FERET still focus on facial recognition challenges (Gates, 2011). Consequently, numerous methods involving essential image handling techniques were used to attain facial properties from images, which were then used to train and identify via distinct classifier algorithms (Kundu *et al*., 2012).

Computer vision operates differently from how humans see. It's a whole system that can find, recognize, process and handle images like humans do (Kafai *et al*., 2014). Critical factors of intellect, such as organization, memory, retrieval, reasoning, estimation and recognition, play significant roles in computer vision (Girshick *et al*., 2016). Early face identification attempts relied on upright images with minimal variations, such as illumination and occlusions. While these techniques were effective for frontal images, they struggled with varied angles or lighting (Roshantharanga *et al*., 2013).

Due to the complexity of facial recognition, many algorithms have been developed for identification, ranging from traditional to advanced methods. Traditional methods focus on identifying and obtaining features or landmarks from the face image, such as the size and position of the nose, eyes, cheekbones and jaw (Bonsor, 2001).

Other methods include compressing facial facts after normalizing a set of facial images, keeping only the necessary details for recognition and comparing this data to a new image (Delac *et al*., 2007). One early successful system used a template identical to point facial features, creating a compressed face representation (Brunelli and Poggio, 1993; Brunelli, 2009).

This study uses a dataset of 280 images from the FEI database, featuring images of twenty individuals across fourteen classes and various positions. The proposed CNN model is designed to be lightweight, with only 16 layers, making it both portable and reliable for real-world applications. We expect this CNN to perform exceptionally well, achieving a training result accuracy of 97.73% and a testing result accuracy of 83.33%. These results will exhibit the model's efficiency in addressing facial recognition challenges while maintaining practicality for real-world use.

Literature Review

Due to rapid technological advancements, computerized facial recognition has become a crucial tool in various fields. These algorithms utilize biometric data from facial features like the shape of the eye socket, nose and chin (Crumpler, 2020; Hiebert *et al*., 2021).

One widely used facial recognition software is FaceMe (Xu *et al*., 2019). Known for its accuracy, speed and flexibility, FaceMe can run on PCs, workstations and servers and integrates with video management systems (VMS). It's often used for security, identifying suspicious individuals in crowds by comparing captured images to those in a database, even if the person is wearing a mask. Additionally, it can detect the body's hotness and send real-time alerts to security administrations (Xu *et al*., 2019). Marriott Hotels have also implemented FaceMe technology to decrease check-in times by around 66.

Apple initiated the iPhone X, featuring Face ID in 1997. This facial recognition system allows users to unlock their phones, buy and use Apple Pay by examining their devices (Komkov and Petiushko, 2021).

Several composite classifiers using K-nearest neighbors were introduced in (Abouelnaga *et al*., 2016), where the authors used PCA (Principal Component Analysis) and KNN alongside CNN to reduce overfitting, increasing accuracy by about 0.7%. Facial recognition accuracy reached 99.95% in 2020 according to a study by the Center of Strategic and International Studies (CSIS).

Such high accuracy has spurred an industrial revolution in facial recognition, enabling it to handle images of varying quality, as noted by the National Institute of Standards and Technology (Grother and Ngan, 2014). However, challenges remain, such as variations in image quality due to different lighting conditions (Crumpler, 2020). Facial recognition is often preferred over other biometric technologies due to its contactless, quick and cost-effective nature.

The demand for facial recognition systems has surged for security in large stores, companies, government buildings and public places, helping identify suspicious individuals (Hill *et al*., 2022). This growth is fueled by the availability of large datasets and the capabilities of Graphics Processing Units (Chen *et al*., 2018).

Figure (1) illustrates the global rise in facial identification usage, with Asia, North America, Europe and other regions contributing. In 2022, the global market was valued at around USD 5.1 billion and is projected to grow at 14.6. The authors in (Sermanet *et al*., 2013) worked on a deep face detector with a deep learning CNN to find faces from different angles without using regression, bounding boxes, or segmentation.

Fig. 1: Facial Recognition market growth

Due to diverse field requirements, researchers are exploring various facial recognition techniques and algorithms to enhance system performance. For instance, (Hammadi *et al*., 2019) focused on multi-view face recognition, while (Hiebert *et al*., 2021) employed a convolutional deep learning neural network on a dataset of 2800 images.

Three classifiers, CNN, K-nearest neighbors and support vector machines, were evaluated on the MNIST dataset (Liu *et al*., 2003). The Multilayer Perceptron (MLP) underperformed in distinguishing digits 9 and 6. Conversely, other classifiers performed well, suggesting that implementing the model using the Keras platform might improve CNN performance. Another study (Abu Ghosh and Maghari, 2017) compared CNN, Deep Belief Networks (DBF) and Deep Neural Networks (DNN) on the MNIST dataset, finding DNN to be the most accurate at 98.08.

To emphasize the significance of feature sets, Chherawala et al. (2016) introduced a weighted recurrent neural network model through weighted votes. The CIFAR-10 dataset (Krizhevsky, 2010) was utilized as a two-layer deep belief network, achieving an accuracy of 78.90% using a GPU unit. The study detailed filter variations and their respective performances, which varied based on the model.

Recent Advancements of CNN

In recent years, CNNs have shown significant advancements in enhanced facial recognition technology performance. These advancements include various CNN architectures, such as depth and efficiency and the integration of innovative layers and practices:

- Deeper networks: The adoption of several deeper architectures, such as VGGNet, res net and Inception (i.e., GoogLeNet), has shown that increasing the depth of the network can significantly improve performance by capturing more complex patterns. In Table (1), we show the depth of many CNN networks. Many research articles reported that these deeper networks provided better-quality accuracy and reliability in facial recognition tasks

- Efficient architectures: New efficient architectures such as MobileNet, ShuffleNet and EfficientNet have been developed to handle many problems in face recognition (Chi *et al*., 2023; Moutsis *et al*., 2023). These architectures can achieve a balanced accuracy and computational efficiency. Thus, it can facilitate the deployment of high-performance facial recognition systems on mobile devices. Efficient Net offers critical advantages over other networks since it can balance accuracy and efficiency using a multiple scaling method that can adjust the network's width, depth and resolution while utilizing few parameters and less computational power
- Attention mechanisms: Incorporating attention mechanisms into the design of CNNs can significantly enhance the implementation of facial recognition systems (Wang, 2022; Daihong *et al*., 2021). A well-known module that can incorporate an attention mechanism is the squeeze and excitation network (SENet). SENet assists the network in concentrating on the highest applicable features by recalibrating channel-wise feature responses. Therefore, the network's competence in distinguishing between refined facial features is better for improving recognition accuracy (Zhang *et al*., 2023)

Further advancement in the abilities and applications of facial recognition technology utilizing CNNs continues to evolve Figs. (2-3).

Table 1: Depth of various CNN architectures

CNN architecture	Depth	References
Inception	22 layers	Anand *et al.* (2020)
ResNet-50	50 layers	Liu (2024)
ResNet-101	101 layers	Ling *et al.* (2019)
ResNet-152	152 layers	Xu and Cloutier (2022)

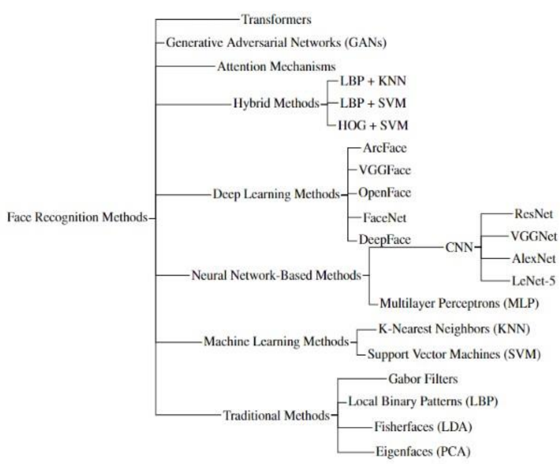

Fig. 2: Overview of face recognition methods

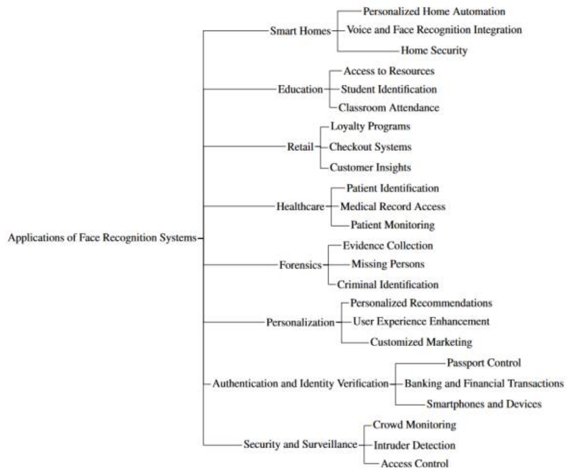

Fig. 3: Applications of face recognition systems

Materials and Methods

This study aimed to design an elegant and consistent CNN for face detection and recognition. This part explains the methodology for addressing the facial detection and recognition problem.

What is Artificial Neural Networks?

ANNs are inspired by the brain's nervous systems (Hsu *et al*., 1995). They mimic the brain's structure, with many neurons working together. Neurons receive inputs through dendrites and send outputs through axons to other neurons via synapses. A neuron activates when input signals surpass a certain threshold, similar to how the human brain functions (Fig. 4) (Mcculloch and Pitts, 1943).

ANNs consist of a group of neurons to optimize input-to-output mathematical functions. These neurons are given a weight with a value that is modified while training to increase performance Fig. (8).

The CNN input layer handles the input images and feeds them to the hidden layer (Nigrin, 1993; Leonard and Kramer, 1990; Jain *et al*., 1996). Several tuning parameters must be chosen before building a CNN. These involve the number of layers, the type of activation function, the filter size and the learning process or algorithm:

$$\phi(S) = \frac{1}{1 + e^{-\sum_{i=0}^{n} w_i x_i}} \tag{1}$$

Preliminary of CNN

CNN is a type of neural network known for its convolutional layers (Huang and Wu, 2017). Typically, CNNs include convolutional and pooling layers, arranged alternately in the network (Krizhevsky *et al*., 2017). CNNs reduce the number of connection weights through weight sharing, creating a simpler architecture as shown in Fig. (5). When processing images, CNNs can handle them directly without complex preprocessing steps. Using

DOI: 10.3844/jcssp.2024.1559.1568

weight sharing, pooling and local receptive fields, CNNs perform well in image transformation tasks like translation, rotation and scaling.

Convolution Layer

CNNs are superior to ANNs since the first layers pick out features that are then passed to a feedforward ANN at the end. Training a CNN takes a long time, but using data pre-processing can speed it up. The convolutional layer generates small-size images that retain many features of the input face (De Bruijn *et al*., 2019). CNN uses several filter sizes to abstract essential features, like blurring enhancement, edge detection and sharpening of objects (Aghdam and Heravi, 2017). The implementation of an image convolution for an image that has dimensions N×M utilizing a filter that has a size of l×k, Equation 2 determines the output feature map size. We define the output width W_{out} and height H_{out}. The following equations provide these definitions:

$$W_{out} = \frac{N - l + 2P_l}{S_k} + 1$$
$$H_{out} = \frac{M - k + 2P_k}{S_k} + 1 \tag{2}$$

The symbols provided in the above equations are presented in Table (2):

$$Y_j^q = \gamma\left(\sum_{I \in I_j} Y_i^{q-1} \times K_{ij}^q + b_j^q\right) \tag{3}$$

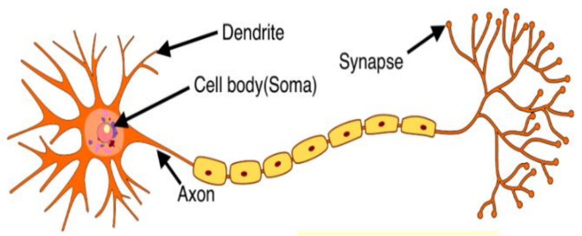

Fig. 4: A biological neuron (Mcculloch and Pitts, 1943)

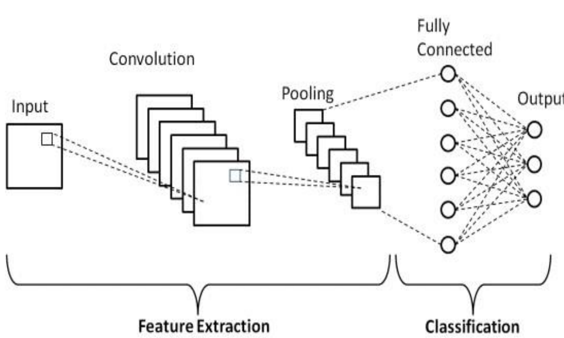

Fig. 5: An overview of a CNN model

Table 2: Definitions of Symbols

Symbol	Description
P_w	Padding for width
P_h	Padding for height
S_w	Strides for width
S_h	Strides for height
q	Layer number
b_j	Bias
Ij	Set of input maps
$\gamma(\cdot)$	Activation function

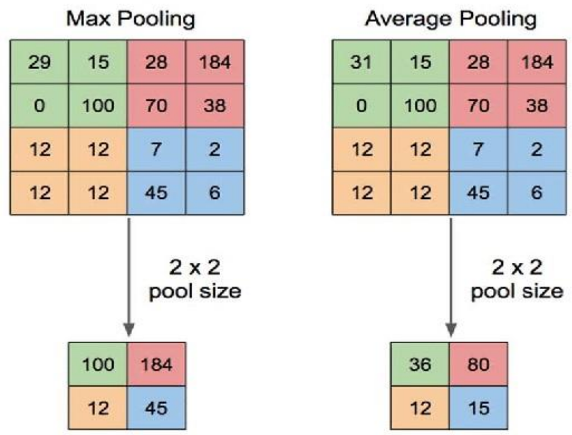

Fig. 6: An example of a max pooling operation

Polling Layer

The pooling layer is an essential component of the CNN structure. It helps aggregate information between various neurons, thus reducing the complexity of the extracted features transferred to the next layers using mathematical formulas like max or average pooling as show in Fig. (6).

These methods help reduce the number of parameters in successive layers. Therefore, the pooling layer helps reduce the CNN model's complexity and softens the effect of overfitting by preventing the network from learning too many details in the training data.

Recited Linear Unit

The adopted activation function is the Rectified Linear Unit that helps add nonlinearity to the inputs, enhancing the learning process. Common activation functions used in the literature include the logistic, tanh, sigmoid and hyperbolic tangent functions. We used the ReLU activation function. ReLU, short for Rectified Linear Unit, returns the input value if it is positive and zero if it is negative. ReLU can achieve faster convergence while training the model (Bhurtel *et al*., 2019). In (Krizhevsky *et al*., 2017), the authors used the ReLU instead of the tanh function. Figure (7), shows the ReLU function and Eq. (4) provides its mathematical expression:

$$R(z) = max(0, z) = \begin{cases} z \, if \; z \le 0 \\ z \, if \; z < 0 \end{cases} \tag{4}$$

DOI: 10.3844/jcssp.2024.1559.1568

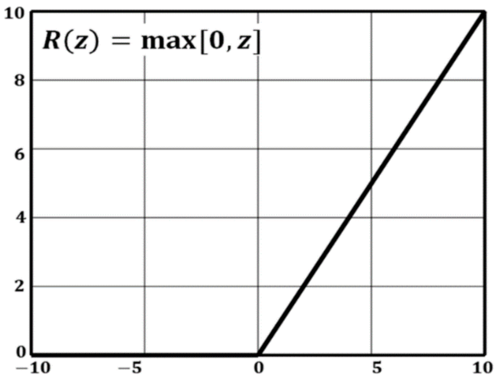

Fig. 7: ReLU: Rectified linear unit

SoftMax Unit

SoftMax is one more activation function frequently used for neurons in the output layer of totally connected networks. It can be described as follows. Table (3) describes the softmax function parameters:

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \tag{5}$$

Fully-Connected Layer

Fully connected layers are extensively employed in various tasks, including image recognition and classification, regression and natural language processing. These layers give substantial flexibility and can learn intricate patterns within the data. However, this flexibility comes at the cost of requiring large amounts of data and computational resources for effective training (Fig. 8 (Bre *et al*., 2018)).

Evaluation Criteria

We calculated recall, accuracy, precision and F1-score. These metrics are standard in classification tasks and are defined as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{9}$$

where, FP, TP, FN and TN represent false positives, true Positives, false negatives and true negatives, respectively.

Proposed CNN-Based Methods

Figure (9) shows the methodology used in this study. It involves designing a CNN model to recognize faces. First, images are collected using a traditional method, like manual capture by a photographer.

Table 3: SoftMax function parameters

Parameter	Description
N	Numbers of Classes
X	Input vector
σ(xi)	Probability of class *i*
Pi	Output probabilities, pi ∈ [0, 1]
$\sum pi$	$\sum_{i=1}^{N} \text{pi} = 1$

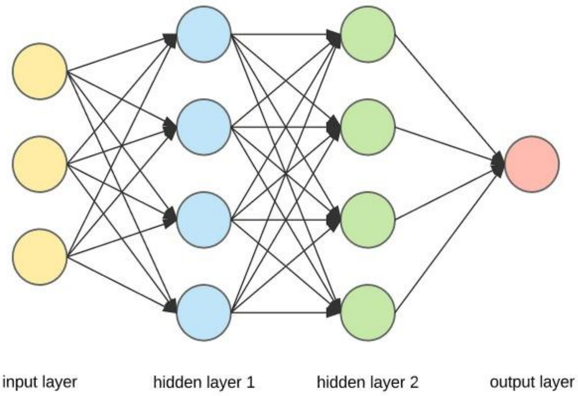

Fig. 8: Fully connected Ann architecture (Bre *et al*., 2018)

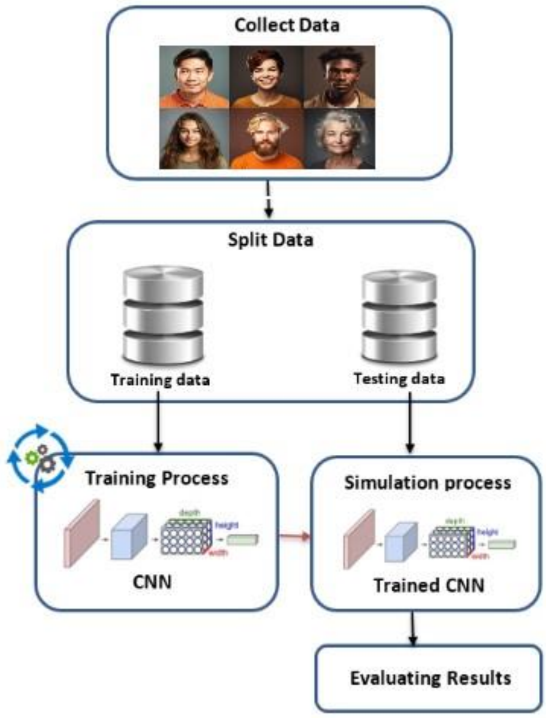

Fig. 9: Proposed CNN-based-detection method (Endri *et al*., 2020)

DOI: 10.3844/jcssp.2024.1559.1568

Fig. 10: Samples images from the FEI face database (De Oliveira Junior and Thomas, 2006)

Dataset

The data set adopted in this research is called the FEI face database (De Oliveira Junior and Thomaz, 2006). It is a Brazilian facial image dataset collected at the AI Laboratory of FEI in São Bernardo do Campo, São Paulo, Brazil. The images were taken between June 2005 and March 2006. The database includes 14 images for each of the 200 individuals, totaling 2800. A homogeneous white background was utilized to capture the photos in color. The images feature individuals in an upright frontal position with profile rotations of up to approximately 180°C. The scale of the images may vary by about 10%, with each original image sized at 640×480 pixels. The dataset mainly contains photos of students and staff at FEI, aged between 19 and 40 years. The data set includes images of 100 males and 100 females. A sample of the FEI database of faces utilized in this experiment is shown in Fig. (10).

Critical Analysis

When implementing research in facial recognition technology, it is critical to report several ethical issues, such as privacy issues and potential biases. These ethical thoughts should guide our steps in the research process and confirm that we are always mindful of the impact of our work.

The utilization of facial images advances significant privacy concerns. It is central to guarantee that all individuals presented in the dataset agree that their images can be used in research, including the purposes for which the photos will be utilized and any possible sharing with third parties or publication of results. Even when permission is granted, extra steps should be considered to anonymize the data to protect the individuals' identities. This may include eliminating or confusing any personal information not necessary for the research.

Protecting the dataset from illegal use and breaches is dominant since facial images include sensitive private data and their exposure could cause identity stealing. Ethical concerns related to the misuse of facial recognition technology can occur in many applications, such as surveillance and law enforcement, which can violate privacy and civil liberties without adequate oversight and accountability. Researchers should study the broader social allegations of their research. It was mentioned in De Oliveira Junior and Thomaz (2006) that the FEI face database is mostly available for research and that researchers are granted permission to use it.

Results and Discussion

The dataset encompasses twenty faces, with fourteen images for each face. The full dataset was used to create the training and testing datasets, with a split of 90% for training and 10% for testing. Figure (11), shows that the intended CNN architecture comprises 16 layers. The proposed CNN's structure and layer descriptions are shown in Fig. (12). Where Fig. (13) displays the results of the facial recognition experiment. The advanced CNN model obtained a training dataset accuracy of 97.73%, demonstrating its ability to identify 97.73% of the samples correctly. On the testing dataset, the accuracy slightly decreases to 83.33%. Figure (17), shows how accuracy and loss change during training. The CNN model achieved a high recognition converges performance of 97%. The implementation of a classification model is rated using a confusion matrix, which provides a comprehensive view of the model's effectiveness Fig. (14). It helps assess the model's accuracy, precision, recall and overall performance.

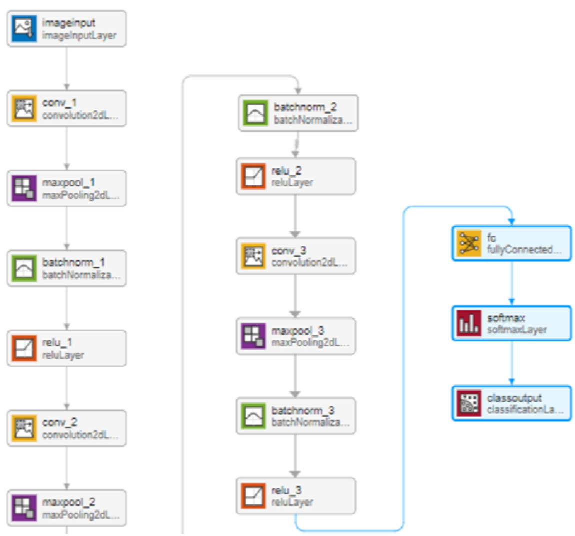

Fig. 11: Proposed CNN model

DOI: 10.3844/jcssp.2024.1559.1568

Layer	Type	Description
'imageinput'	Image Input	227×227×3 images with 'zerocenter' normalization
'conv_1'	Convolution	8 filters, 5×5×3 convolutions with stride [1 1] and padding 'same'
'maxpool_1'	Max Pooling	2×2 max pooling with stride [2 2] and padding [0 0 0 0]
'batchnorm_1'	Batch Normalization	Batch normalization with 8 channels
'relu_1'	ReLU	ReLU activation function
'conv_2'	Convolution	16 filters, 5×5×8 convolutions with stride [1 1] and padding 'same'
'maxpool_2'	Max Pooling	2×2 max pooling with stride [2 2] and padding [0 0 0 0]
'batchnorm_2'	Batch Normalization	Batch normalization with 16 channels
'relu_2'	ReLU	ReLU activation function
'conv_3'	Convolution	32 filters, 5×5×16 convolutions with stride [1 1] and padding 'same'
'maxpool_3'	Max Pooling	2×2 max pooling with stride [2 2] and padding [0 0 0 0]
'batchnorm_3'	Batch Normalization	Batch normalization with 32 channels
'relu_3'	ReLU	ReLU activation function
'fc'	Fully Connected	20 fully connected layer
'softmax'	Softmax	Softmax activation function
'classoutput'	Classification Output	Cross-entropy loss with '1' and 19 other classes

Fig. 12: Description of the CNN layers

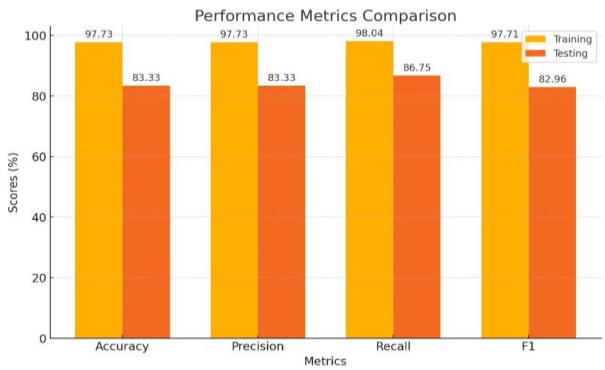

Fig. 13: Performance matrices comparison

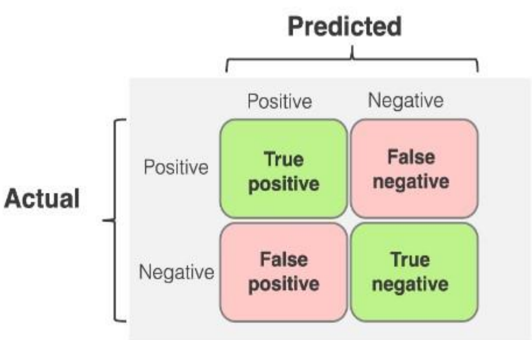

Fig. 14: Confusion matrix

The main characteristics of a confusion matrix include:

- True Positives (TP): The cases where the defined model properly predicts the positive class
- True Negatives (TN): The cases where the defined model properly predicts the negative class
- False Positives (FP): The cases where the defined model improperly predicts the positive class (Type I error)
- False Negatives (FN): The cases where the defined model improperly predicts the negative class (Type II error)

The confusion matrices developed are shown in Figs. (15-16).

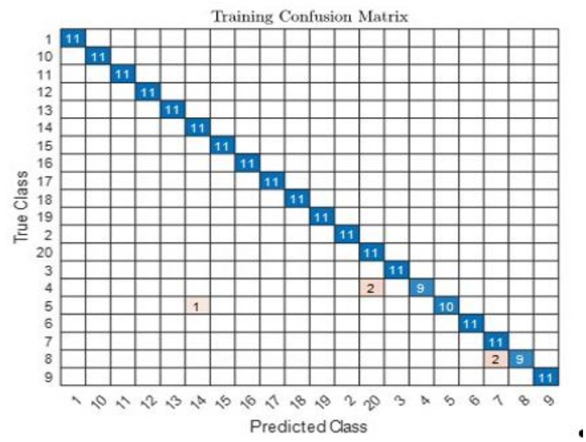

Fig. 15: Training case: Confusion matrix

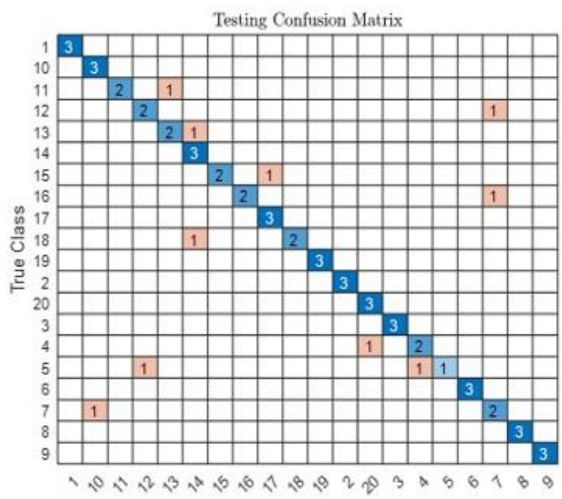

Fig. 16: testing case: Confusion matrix

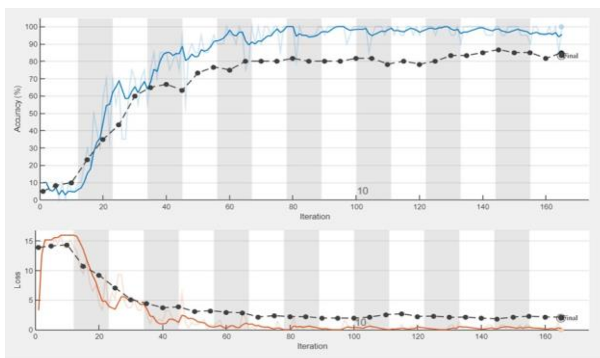

Fig. 17: Accuracy and loss of the developed CNN model

Future Research Directions

Facial recognition technology has evolved considerably, from simple image processing methods to complex deep learning models. These systems have been extended to cover various applications, such as security, authentication, healthcare, retail, education and smart homes. The field of research has started with conventional or traditional methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) Zhao *et al.* (2002); Shailaja and Patil (2014) to cutting-edge deep learning methods, including CNNs, GANs and transformers Guo and Zhang (2019); Moutsis *et al.* (2023). These advancements resulted in better accuracy, reliability and easy deployment of the facial recognition models.

Future innovations could include integrating facial recognition into augmented reality (AR) and virtual reality (VR) systems. Thus, we could explore the extra domains of applications and augmented user experiences. The applicability of these types of systems to real-time applications will lead to the design of facial recognition systems that are more reasonable for large-scale applications. Many challenges are still involved in the design of such systems, such as data security, encryption and secure storage for data protection. Although facial recognition technology presents significant potential advantages in various applications, future research will explore legal, ethical and technical challenges.

Conclusion

This research presented a novel method by developing a lightweight CNN architecture exclusively designed for facial recognition. With its 16 layers, our proposed CNN was trained on a dataset of 280 images representing 20 faces with varying orientations. The model's performance was extraordinary, attaining a training accuracy of 97.73% and a testing accuracy of 83.33%. Our results showed that the proposed CNN handles the facial recognition challenges extremely well, outperforming many existing methods in accuracy and efficiency. The proposed model is simple, applicable for mobile applications, and can be uploaded to low-computation devices. This reliable CNN architecture suggested the advantage of the proposed CNN model compared to complex models presented in the literature.

Acknowledgment

The research reported in this publication was supported by the Deanship of Scientific Research and Innovation at Al-Balqa Applied University in Jordan.

Funding Information

This research is a self-funding no external funding.

Ethics

The authors emphasize that the research article offered in this publication is original and has not been previously published in any other journal.

References

Abouelnaga, Y., Ali, O. S., Rady, H., & Moustafa, M. (2016). CIFAR-10: KNN-Based Ensemble of Classifiers. *2016 International Conference on Computational Science and Computational Intelligence (CSCI)*, 1192–1195. https://doi.org/10.1109/csci.2016.0225

Abu Ghosh, M. M., & Maghari, A. Y. (2017). A Comparative Study on Handwriting Digit Recognition Using Neural Networks. *2017 International Conference on Promising Electronic Technologies (ICPET)*, 77–81. https://doi.org/10.1109/icpet.2017.20

Aghdam, H. H. and Heravi, E. J. (2017). Guide to Convolutional Neural Networks: A Practical Application to Traffic-Sign Detection and Classification. Springer Publishing Company, Incorporated, 1st edition.

Anand, R., Shanthi, T., Nithish, M. S., & Lakshman, S. (2020). *Face Recognition and Classification Using GoogleNET Architecture* (K. Nath Das, J. Chand Bansal, K. Deep, A. K. Nagar, Ponnambalam Pathipooranam, & C. Naidu, Eds.; 1st ed., Vol. 1, pp. 261–269). Springer Singapore. https://doi.org/10.1007/978-981-15-0035-0_20

Bhurtel, M., Shrestha, J., Lama, N., Bhattarai, S., Uprety, A., & Guragain, M. K. (2019). Deep Learning Based Seed Quality Tester. *Science, Engineering and Technology (SET) Conference – 2019*, 1–10.

Bledsoe, W. W. (1964). *The model method in facial recognition Technical report PRI-15*. Panoramic Research, Inc.

Bledsoe, W. W. (1966a). *Man-Machine Facial Recognition: Report on a Large-scale Experiment*. Panoramic Research, Inc.

Bledsoe, W. W. (1966b). Some Results on Multicategory Pattern Recognition. *Journal of the ACM*, *13*(2), 304–316. https://doi.org/10.1145/321328.321340

Bledsoe, W. W. (1968). *Semiautomatic Facial Recognition*. Stanford Research Institute, Menlo Park.

Bledsoe, W. W., & Chan, H. (1965). *A man-Machine Facial Recognition System - Some Preliminary Results*. Panoramic Research, Inc.

Bonsor, K., Johnson, R., & Taras, Z. (2001). *How Facial Recognition Systems Work*. HowStuffWorks. https://electronics.howstuffworks.com/gadgets/high-tech-gadgets/facial-recognition.htm

Bower, G. H., & Karlin, M. B. (1974). Depth of Processing Pictures of Faces and Recognition Memory. *Journal of Experimental Psychology*, *103*(4), 751–757. https://doi.org/10.1037/h0037190

Bre, F., Gimenez, J. M., & Fachinotti, V. D. (2018). Prediction of Wind Pressure Coefficients on Building Surfaces Using Artificial Neural Networks. *Energy and Buildings*, *158*, 1429–1441. https://doi.org/10.1016/j.enbuild.2017.11.045

Brunelli, R. (2009). Template Matching Techniques in Computer Vision: Theory and Practice. Wiley.

Brunelli, R., & Poggio, T. (1993). Face Recognition: Features versus Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*(10), 1042–1052. https://doi.org/10.1109/34.254061

Chen, J.-C., Ranjan, R., Sankaranarayanan, S., Kumar, A., Chen, C.-H., Patel, V. M., Castillo, C. D., & Chellappa, R. (2018). Unconstrained Still/Video-Based Face Verification with Deep Convolutional Neural Networks. *International Journal of Computer Vision*, *126*(2–4), 272–291. https://doi.org/10.1007/s11263-017-1029-3

Chherawala, Y., Roy, P. P., & Cheriet, M. (2016). Feature Set Evaluation for Offline Handwriting Recognition Systems: Application to the Recurrent Neural Network Model. *IEEE Transactions on Cybernetics*, *46*(12), 2825–2836. https://doi.org/10.1109/tcyb.2015.2490165

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), 4–19. https://doi.org/10.3991/ijim.v17i23.40867

Crumpler, W. (2020). How accurate are facial recognition systems–and why does it matter. *Center for Strategic and International Studies*, *14*.

Daihong, J., Lei, D., & Jin, P. (2021). Facial expression recognition based on attention mechanism. *Scientific Programming*, *2021*(1), 6624251. https://doi.org/10.1155/2021/6624251

De Bruijn, A., Muhonen, V., Albinonistraat, T., Fokkink, W., Bloem, P., and Analytics, B. (2019). Detecting offensive language using transfer learning.

Delac, K., Grgic, M., & Grgic, S. (2007). Image Compression Effects in Face Recognition Systems. In *Face Recognition*. IntechOpen. https://doi.org/10.5772/4832

De Oliveira Junior, L. and Thomaz, C. E. (2006). Captura e alinhamento de imagens: Um banco de faces brasileiro. Technical report, Department of Electrical Engineering, FEI.

Endri, E., Sheta, A., & Turabieh, H. (2020). Road Damage Detection Utilizing Convolution Neural Network and Principal Component Analysis. *International Journal of Advanced Computer Science and Applications*, *11*(6). https://doi.org/10.14569/ijacsa.2020.0110682

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2016). Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*(1), 142–158. https://doi.org/10.1109/tpami.2015.2437384

Grother, P., & Ngan, M. (2014). *Face Recognition Vendor Test (FRVT)*. National Institute of Standards and Technology. https://doi.org/10.6028/nist.ir.8009

Guo, G., & Zhang, N. (2019). A Survey on Deep Learning Based Face Recognition. *Computer Vision and Image Understanding*, *189*, 102805. https://doi.org/10.1016/j.cviu.2019.102805

Hammadi, O. I., Dawah Abas, A., & Hammad Ayed, K. (2019). Face recognition using deep learning methods a review. *International Journal of Engineering & Technology*, *7*(4), 6181–6188. https://doi.org/10.14419/ijet.v7i4.22375

Hiebert, J., Mazhar, F., Derosa, M., & Sheta, A. (2021). Facial Recognition Using a Lightweight Deep Neural Networks. *Journal of Advanced Computer Science & Technology*, *10*(1), 1–8. https://doi.org/10.14419/jacst.v10i1.31632

Hill, D., O'Connor, C. D., & Slane, A. (2022). Police Use of Facial Recognition Technology: The Potential for Engaging the Public through Co-Constructed Policy-Making. *International Journal of Police Science & Management*, *24*(3), 325–335. https://doi.org/10.1177/14613557221089558

Hsu, K., Gupta, H. V., & Sorooshian, S. (1995). Artificial Neural Network Modeling of the Rainfall-Runoff Process. *Water Resources Research*, *31*(10), 2517–2530. https://doi.org/10.1029/95wr01955

Huang, W., & Wu, Q. (2017). Image Retrieval Algorithm Based on Convolutional Neural Network. In *Current Trends in Computer Science and Mechanical Automation* (Vol. 1, pp. 304–314). De Gruyter Open. https://doi.org/10.1515/9783110584974-034

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial Neural Networks: A Tutorial. *Computer*, *29*(3), 31–44. https://doi.org/10.1109/2.485891

Kafai, M., An, L., & Bhanu, B. (2014). Reference Face Graph for Face Recognition. *IEEE Transactions on Information Forensics and Security*, *9*(12), 2132–2143. https://doi.org/10.1109/tifs.2014.2359548

Gates, K. A. (2011). *Our biometric future: Facial recognition technology and the culture of surveillance* (Vol. 2). NYU Press. ISBN-10: 9780814732793.

Komkov, S., & Petiushko, A. (2021). AdvHat: Real-World Adversarial Attack on ArcFace Face ID System. *2020 25th International Conference on Pattern Recognition (ICPR)*, 819–826. https://doi.org/10.1109/icpr48806.2021.9412236

Krizhevsky, A. (2010). Convolutional Deep Belief Networks on Cifar-10. *Unpublished Manuscript*, *40*(7), 1–9.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90. https://doi.org/10.1145/3065386

Kundu, M. K., Mitra, S., Mazumdar, D., & Pal, S. K. (Eds.). (2012). *Perception and Machine Intelligence* (1st Ed.). Springer. https://doi.org/10.1007/978-3-642-27387-2

Leonard, J., & Kramer, M. A. (1990). Improvement of the backpropagation algorithm for training neural networks. *Computers & Chemical Engineering*, *14*(3), 337–341. https://doi.org/10.1016/0098-1354(90)87070-6

Ling, H., Wu, J., Wu, L., Huang, J., Chen, J., & Li, P. (2019). Self-Residual Attention Network for Deep Face Recognition. *IEEE Access*, *7*, 55159–55168. https://doi.org/10.1109/access.2019.2913205

Liu, C.-L., Nakashima, K., Sako, H., & Fujisawa, H. (2003). Handwritten Digit Recognition: Benchmarking of State-of-the-Art Techniques. *Pattern Recognition*, *36*(10), 2271–2285. https://doi.org/10.1016/s0031-3203(03)00085-2

Liu, J. (2024). Face Recognition Technology Based on Resnet50. *Applied and Computational Engineering*, *39*, 160–165.

McCulloch, W. S., & Pitts, W. (1943). A logical Calculus of the Ideas Immanent in Nervous Activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115–133. https://doi.org/10.1007/bf02478259

Moutsis, S. N., Tsintotas, K. A., Kansizoglou, I., & Gasteratos, A. (2023). Evaluating the Performance of Mobile-Convolutional Neural Networks for Spatial and Temporal Human Action Recognition Analysis. *Robotics*, *12*(6), 167. https://doi.org/10.3390/robotics12060167

Nigrin, A. (1993). *Neural Networks for Pattern Recognition. A Bradford Book*. The MIT Press.

Nils, J. N., (2009). *The Quest for Artificial Intelligence*. Cambridge University Press. https://doi.org/10.1017/cbo9780511819346

Roshantharanga, J. G., Samarakoon, S. M. S. C., Karunarathne, T. A. P., Liyanage, K. L. P. M., Gamage, M. P. A. W., & Perera, D. (2013). Smart Attendance Using Real-Time Face Recognition (smart - fr). *SAITM Research Symposium on Engineering Advancements*, 41–44.

Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & Lecun, Y. (2013). Overfeat: Integrated Recognition, Localization, and Detection Using Convolutional Networks. *ArXiv*, arXiv:1312.6229.

Shailaja A., & Patil, P. J. D. (2014). Principle component analysis (pca) and linear discriminant analysis (lda) based face recognition. National Conference on advances in Communication and Computing, CACC2014 (3):1–5.

Wang, Q. (2022). *Face Recognition with Attention Mechanisms*.

Xu, W., & Cloutier, R. S. (2022). A Facial Expression Recognizer using Modified ResNet-152. *EAI Endorsed Transactions on Internet of Things*, *7*(28), e5. https://doi.org/10.4108/eetiot.v7i28.685

Xu, H., Su, X., Wang, Y., Cai, H., Cui, K., & Chen, X. (2019). Automatic Bridge Crack Detection Using a Convolutional Neural Network. *Applied Sciences*, *9*(14), 2867. https://doi.org/10.3390/app9142867

Zhang, M., Zhang, Y., & Zhang, Q. (2023). Attention-Mechanism-Based Models for Unconstrained Face Recognition with Mask Occlusion. *Electronics*, *12*(18), 3916. https://doi.org/10.3390/electronics12183916

Zhao, W., Chellappa, R., & Krishnaswamy, A. (2002). Discriminant analysis of principal components for face recognition. *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, 73–85. https://doi.org/10.1109/afgr.1998.670971