[1]Chennakesh S,
2Santosh Gayakwad,
[3]Rupesh Gupta,
[4]Sayyad Jilani,
[5]Dr. Pushpa Mamoria,
[6]Rahul Vudatha

Deep Learning For Face Recognition: Ai-Powered Solutions With Convolutional Neural Networks

Journal of Electrical Systems

Abstract: - Face recognition is becoming more prevalent in addressing several social issues, such as enhancing personal security and verifying identity. Facial recognition is a type of biometric technology that is employed with other commonly used biometric applications, like iris recognition, and vein pattern recognition, along with fingerprint recognition. This type of recognition is a process that identifies an individual by analysing specific characteristics of their physical features. Deep Learning (DL) is a subset of machine learning (ML) that specializes in using neural networks to process images and identify patterns. One of its many applications is in face recognition. DL has led to the widespread use of Convolution Neural Network (CNN) grounded facial recognition technology, making it the leading approach in face identification. This investigation's chief goal is to explore DL for Face Recognition, which involves the use of Artificial Intelligence (AI) - Powered Solutions with Convolutional Neural Networks. A methodical strategy is suggested to optimize the parameters and improve the system's performance. CNNs have similarities with regular neural networks, yet they specifically assume the inputs happe to be images. This allows designers to integrate certain qualities into the architecture. This study offered the construction and gauging of a real-time facial recognition system utilizing CNN Architecture. The offered system and also CNN architecture are assessed by optimizing several parameters of the CNN in order to boost the recognition correctness of the designed system. So, the offered system achieved a maximum recognition correctness of 98.75% when utilising Kaggle databases for inputs and also 98.00% when utilising real-time inputs.

Keywords: DL; Face Recognition; CNN Architecture; AI; Accuracy; Real – time Facial Recognition.

INTRODUCTION

Face recognition is a technique that is employed to determine or authenticate the identification of a person by analysing their facial features. Face recognition is utilized in diverse applications, counting an automated classroom attendance management system [1], access-restricted areas' surveillance, such as living spaces, or for intruder detection [2], identification of celebrities inside public spaces [3], and recognition of household occupants by networked home automation systems [4], among others. The majority of face recognition systems (FCSs) are comprised of two primary modules: feature extraction along with classifier. FCSs have utilized different combinations of feature extraction and also classifier techniques.

[1] Research Scholar, Capital University, Jharkhand, India
Email ID: chennakesh.2015@gmail.com
[2]Ex-Visiting Professor, BITS, Pilani, MCA, MSc (Psychology), MTech (Software Systems), MPhil(Computer Science), PGP (IIM, Calcutta), PGP (Purdue University)
Email ID: iksantosh01@gmail.com
[3]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
Email ID: rupesh.gupta@chitkara.edu.in
[4]M.H. Saboo Siddik College of Engineering, Mumbai, India
Email ID: jilani.sayyad@gmail.com
[5]Senior Lecturer State and city KANPUR (U.P.)
Department of Computer Applications
Chhtrapati Shahu Ji Maharaj University KANPUR (U.P.)
Email ID: p.mat76@gmail.com
[6]Associate Engineer (Working professional)
Vellore Institute of Technology, Andhra Pradesh, India
Email ID: rahulvudatha@gmail.com

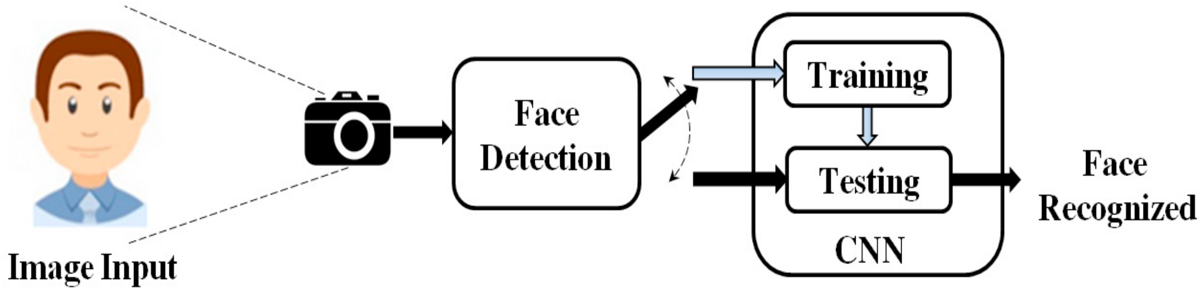

Figure 1: Real-time face recognition block diagram [5]

In recent years, there have been notable breakthroughs in face recognition technology, primarily driven by the progress made in deep learning techniques, specifically CNNs. These AI-driven solutions have transformed the industry by offering extremely precise and effective techniques for detecting and authenticating human faces [3]. The fundamental basis of deep learning for face recognition resides in its capacity to autonomously acquire hierarchical representations of facial characteristics from extensive datasets. CNNs outperform traditional methods that relied on manually designed features and faced difficulties in handling fluctuations in lighting, position, and emotion. CNNs are highly effective in capturing complex patterns by utilizing several layers of abstraction. The initial layers inside a CNN are accountable for spotting basic features like edges and textures. As move deeper into the network, the subsequent levels become capable of capturing more intricate structures such as facial landmarks. Ultimately, these deeper layers are able to identify unique characteristics that distinguish one face from another [6].

The hierarchical learning mechanism allows CNNs to accurately and reliably identify faces in various situations. CNNs commonly consist of layers like convolutional, pooling, and fully connected layers, each of which has a crucial function in extracting features and performing classification [5]. The subsequent section provides an in-depth analysis of previous literature pertaining to this study.

LITERATURE REVIEW

The table below provides an in-depth analysis of prior investigation on Deep Learning for Face Recognition, specifically focusing on AI-Powered Solutions with Convolutional Neural Networks.

AUTHORS AND YEARS	METHODOLOGY	FINDINGS
Mehendale (2020) [7]	This study introduced facial emotion recognition using convolutional neural networks (FERC). The FERC uses a two-part CNN to eliminate background and retrieve facial feature vectors.	The FERC method excels in working with orientations under 30° due to its unique 24-digit long EV feature matrix. Background removal greatly improved the accuracy of identifying emotions.
Abdullah and Abdulazeez (2021) [8]	Provided a brief introduction of FER fields of application and publically accessible databases and studied the latest and current FER reviews utilizing CNN algorithms.	Finally, everyone obtained good results, especially accuracy, with various rates and data sets, which affects results.
Vu et al., (2022) [9]	This study used RetinaFace, a joint extra-supervised and self-supervised multi-task learning face detector that can handle various face scales, as a fast yet effective encoder to	With an 87% f1-score on the COMASK20 dataset and 98% on the Essex dataset, our suggested system outperformed Dlib and InsightFace, proving its efficacy and adaptability.

	recognize the masked face using deep learning and Local Binary Pattern (LBP) features.	
Canedo et al., (2023) [10]	This study suggested image processing techniques for validating new personal photos, including face detection, recognition, cropping, image quality assessment, head posture estimation, gaze estimation, blink detection, and sunglasses detection. These algorithms verify submitted photos based on established criteria.	Results for individual image processing algorithms range from 92% to 100% accuracy, depending on the algorithm being tested.
Pinzón-Arenas et al., (2024) [11]	A study analysed eight convolutional architecture base models using transfer of learning, as well as two proposed models, shallow CNN and shallow DAG-CNN, which had six convolution layers.	GoogLeNet and ResNet-101 were shown to be the most effective networks for this application, accurately recognizing two users and discriminating against non-database subjects.

Research Gap

Although there have been notable improvements, there is still a lack of study in addressing the bias and fairness of deep learning models used for face recognition. These algorithms frequently demonstrate unequal performance among various demographic groups. Moreover, there is a requirement for enhanced comprehensibility and openness in CNN -based facial recognition systems in order to cultivate trust and achieve broader approval. Moreover, the task of effectively training and implementing these models on devices with limited resources, while maintaining accuracy, is a continuous subject of research. The chief focus of this study is to investigate DL for Face Recognition using AI-Powered Solutions using Convolutional Neural Networks.

METHODOLOGY

The performance gauging of the offered FCS was conducted utilising the regular Kaggle database[2], which consists of 10 photographs from 40 persons, resulting in a total of 400 images. The figure below displays the samples of 40 persons from the Kaggle database. Out of a total of 400 photos, 320 photographs (8 images from each of the 40 participants) were allocated for training, yet the residual 80 images were then set aside for testing.

[2] https://www.kaggle.com/datasets/kasikrit/att-database-of-faces

Figure 2: Samples from Kaggle database[1]

CNN architectures exhibit variability across designers, with the arrangement of layers' subject to modification through recurrent evaluations to achieve optimal recognition correctness. The offered work utilizes the CNN architecture depicted in the illustration below. After assessing several permutations of sequence layers. Also, the CNN architecture being presented is implemented using Keras, an Open Source Neural Network toolkit, which runs on top of Tensorflow. So, the convolutional layer consists of both convolutional and rectified linear unit (RELU) layers. Also, the input image recorded from the camera is initially processed by the Viola Jones algorithm for the purpose of face detection. And, the facial image is first cropped and then converted to grayscale. It is then enlarged to 120x120 pixels and sent through the initial convolution layer, consisting of 32 filters with dimensions of 3x3 pixels. It is important to mention that said filters' weights are initially assigned random values and they are adjusted utilising the backpropagation technique across a few epochs to obtain the final weights for these filters. The weights obtained at the end of the process are then utilized in the categorization stage. The result of the initial convolutional layer, which consists of 32 filters as previously indicated, is then passed as input to the second convolutional layer. The second layer utilises a separate set of 32 filters, each with a size of 3x3 pixels, to produce an output. The output from the second convolutional and rectified linear unit (CONV+RELU) layer is passed to a pooling layer using the max pooling function, with a window size of 4x4 pixels. A depiction of the output of a POOL layer employing both max pooling along with average pooling. During the evaluation, it was noticed that max pooling yielded higher accuracy compared to average pooling for the suggested work. Therefore, max pooling was chosen for implementation in this study.

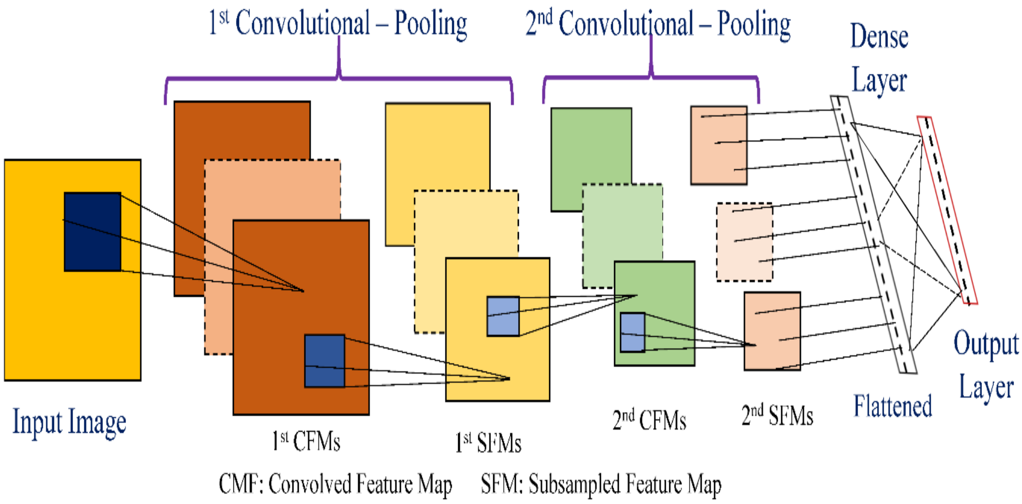

Figure 3: CNN Architecture for the proposed study

RESULTS AND DISCUSSIONS

The suggested system's performance evaluation is conducted by altering the count of filters in the convolution layer along with the convolution filter's window size, while using different pooling window sizes. The assessment findings, as well as the system's identification accuracy, are graphed. The x-axis represents the convolution filter's window size, whereas the y-axis denotes the count of filters in the convolution layer. It has been noted that the use of a convolution filter measuring 3x3 pixels having 32 filters resulted in a maximum recognition correctness of 98.75% for the suggested system. This accuracy was achieved by employing a pooling window size of 2x2 and 4x4 pixels. The proposed technique and CNN architecture are comparable to the work documented in the literature. The suggested work achieves an improvement in recognition accuracy by optimizing the count of convolution filters, convolution filters' window size, along with pooling.

Following the successful gauging and testing of the offered system utilising a typical Kaggle dataset, the system performance is assessed for real-time inputs captured by a camera. The evaluation of the offered real-time system is based on samples from 5 persons. A total of 200 photos are recorded, with 40 images taken of each participant. To determine the recognition correctness of the offered system for real-time input, a total of 200 photographs were employed. Out of these, 100 images (being 20 images directly from 5 individuals) were utilised for training, while the residual 100 images were used for testing. The experiments were conducted to determine the optimal count of convolution filters alongside the window size for the convolution along with pooling layers in real-time systems. The assessment results show the relationship between the window size for the convolution filter (x-axis) and the count of filters in the convolution layer (y-axis). The real-time system achieved a maximum recognition correctness of 98.00% by utilizing 32 convolution filters having pooling window sizes of 2x2, 3x3, and 4x4 pixels, as well as varying window sizes for the convolution filter. So, the output outcomes achieved throughout the live demonstration of the offered real-time FCS are exhibited in a snapshot. The system takes individual photos as inputs and displays their identities straight on the top left. The suggested method initially identifies a face inside the image and subsequently distinguishes the face, displaying the person's identification.

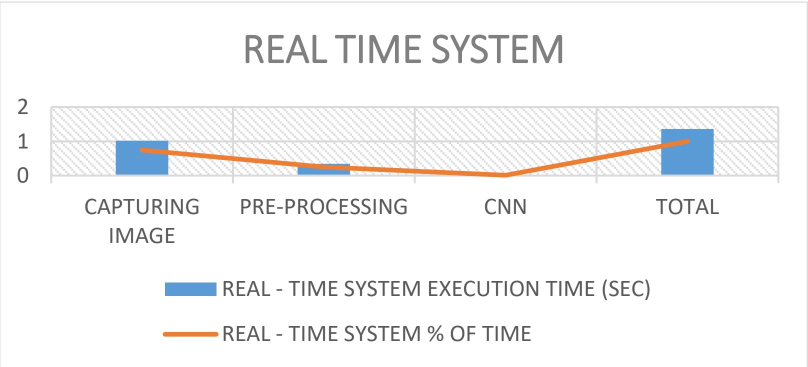

Figure 4: Results of Real – time System

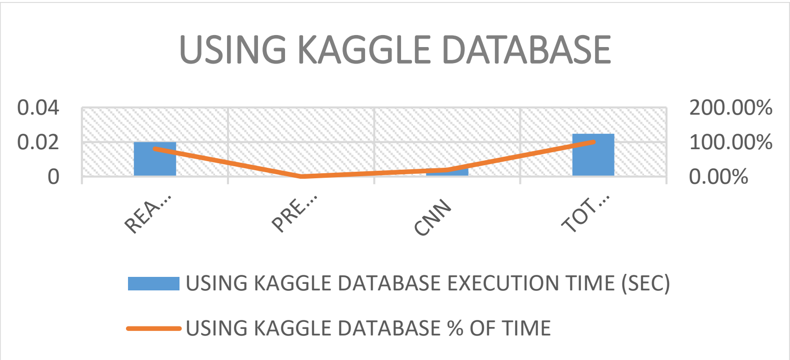

Figure 5: Results of Using Kaggle database

The analysis of deep learning performance for face recognition using CNNs demonstrates clear disparities between real-time system implementations and database-driven approaches, as seen by the presented execution time statistics. The image capture process in a real-time system accounts for 74.07% of the execution time, indicating its high level of intensity. The pre-processing stage accounts for 24.93% of the overall time and involves tasks such as scaling, normalization, and augmentation to prepare the image for the CNN. The CNN design, although crucial for face recognition, is remarkably efficient, accounting for only 1.00% of the whole execution time, equivalent to 0.0136 seconds. This highlights the optimized and rapid performance of contemporary CNN architectures in managing the recognition process. On the other hand, when utilizing a pre-existing dataset such as the Kaggle database, the way the execution works changes considerably. The dataset files reading time accounts for 80.06% of the total execution time, indicating the additional time needed to access and load data from storage. In contrast to the real-time system, the pre-processing time is minimal or non-existent, maybe because the dataset already contains pre-processed and pre-formatted data. The execution time of the CNN, while still very short, accounts for a significant amount of the total duration at 19.36%, which is equivalent to 0.0048 seconds. The higher percentage gain in comparison to the real-time scenario indicates that the efficiency of the CNN remains constant. However, the significantly lower total execution time of 0.0248 seconds highlights its more prominent role.

Face recognition technology has undergone a revolution due to deep learning, namely CNNs, which offer incredibly accurate and effective AI-powered solutions. CNNs are perfect for facial recognition tasks because they are excellent at automatically deriving the spatial hierarchies of features from photos. These networks are capable of identifying and extracting complex patterns and information from facial photos, including forms, edges, and textures, all of which are essential for differentiating between faces. AI facial recognition is being used in a variety of practical applications, from social media and smartphone personalization to access control and security systems. Deep learning and CNN integration for face recognition not only boosts these systems' dependability and performance but also creates new opportunities for innovation across a range of industries.

CONCLUSION

The offered system along with CNN architecture are assessed by adjusting various parameters of the CNN in order to improve the recognition correctness of the designed system. The submitted system attains a maximum recognition correctness of 98.75% when using Kaggle inputs and 98.00% when utilizing real-time inputs. So, the submitted work can be readily customized for many consumer applications, including face detection-grounded home automation, and device control, and attendance system, and intrusion detection. The study highlights the different computational requirements and effectiveness of facial recognition systems based on the operating situation. Real-time systems are primarily influenced by the practical aspects of capturing and processing images, which require improvements in order to minimize delay. However, database-driven systems prioritize efficient data retrieval mechanisms to reduce the time spent on reading operations, even though they are designed for recognition tasks. Both situations demonstrate the resilience and adaptability of CNNs in face recognition applications, since they are able to retain high performance across various stages of the process. These observations can provide guidance for future enhancements and optimizations, whether the goal is to speed up the process of capturing real-time images or improve the efficiency of handling data in existing datasets.

REFERENCES

[1] R. Samet and M. Tanriverdi, "Face Recognition based Mobile Automatic Classroom Attendance Management System," in *International Conference on Cyberworlds*, Chester, United Kingdom: IEEE Computer Society, 2017, pp. 253–256.

[2] P. Fahad, M. Mahmudul, M. Susan, and M. Moslehuddin, "Face recognition based real time system for surveillance," in *Intelligent Decision Technologies*, vol. 11, IOS Press, 2017, pp. 79–92.

[3] H. Ouanan, M. Ouanan, and B. Aksasse, “Pubface: Celebrity face identification based on deep learning,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 353, p. 012022, 2018.

[4] N. Satani, Information Technology, Parul University, Vadodara, India., S. Patel, S. Patel, Information Technology, Parul University, Vadodara, India., and Information Technology, Parul University, Vadodara, India., “AI powered glasses for visually impaired person,” *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 9, no. 2, pp. 316–321, 2020.

[5] Pranav and Manikandan, “Design and evaluation of a real-time face recognition system using convolutional neural networks,” *Procedia Comput. Sci.*, vol. 171, pp. 1651–1659, 2020.

[6] S. Almabdy and L. Elrefaei, “Deep convolutional neural network-based approaches for face recognition,” *Appl. Sci. (Basel)*, vol. 9, no. 20, p. 4397, 2019.

[7] N. Mehendale, “Facial emotion recognition using convolutional neural networks (FERC),” *SN Appl. Sci.*, vol. 2, no. 3, 2020.

[8] S. M. S. Abdullah and A. M. Abdulazeez, “Facial expression recognition based on deep learning convolution neural network: A review,” *Journal of Soft Computing and Data Mining*, vol. 2, no. 1, pp. 53–65, 2021.

[9] H. N. Vu, M. H. Nguyen, and C. Pham, “Masked face recognition with convolutional neural networks and local binary patterns,” *Appl. Intell.*, vol. 52, no. 5, pp. 5497–5512, 2022.

[10] D. Canedo, J. Vieira, A. Gonçalves, and A. J. Neves, “AI-Powered Management of Identity Photos for Institutional Staff Directories,” in *VISIGRAPP (4: VISAPP)*, 2023, pp. 804–811.

[11] J. O. Pinzón-Arenas, R. Jimenez-Moreno, and J. E. Martinez Baquero, “Comparison of convolutional neural network models for user’s facial recognition,” *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 14, no. 1, p. 192, 2024.