International Journal of Advanced Trends in Multidisciplinary Research
ISSN: 3139-0986, Volume 2 Issue 7
July 2026, PP 1-5

Deep Learning-Based Face Recognition for Automated Attendance Management

[1]M. Radhika, [2]M.Prasanna Kumar, [3]M. Sivaparvathi
[1,2]Assoc. Prof, Department of CSE, MAM Women's Engineering College, Narasaraopet, Palnadu, A.P., India.
[3]Asst. Prof., Department of CSE, MAM Women's Engineering College, Narasaraopet, Palnadu, A.P., India.

***Abstract* – The internet and healthcare industries rely heavily on deep learning and transfer learning. Almost every sector of the modern internet economy is dependent on face recognition software. The environment is always evolving, and so are automated face recognition (AFR) systems. With Smart Attendance with Real-Time Face Recognition, keeping track of students' daily attendance is a breeze. One approach that has been suggested builds a dataset using images of students in class. We employ the Multitask Convolution Neural Network technique to detect and identify faces. Facial recognition will be followed by data submission. The suggested approach was contrasted with other alternatives, including the Haar Cascade Convolution Neural Network Classification method. The suggested approach outperformed the alternatives in terms of results.**

***Keywords*- Face Detection, Convolutional Neural Networks, and Student Attendance**

I. INTRODUCTION

Facial recognition technology may also be used to monitor who shows up at other establishments including schools, corporations, and colleges. The development of a facial recognition-based attendance system is one potential use of this technology. The present manual attendance method is quite labour-intensive to maintain. Furthermore, it could be necessary to use a proxy. You should be aware that proxies might appear. For the sake of the school's record keeping, students are expected to preserve accurate attendance records. Each and every one of these establishments accomplishes this in its own unique way. Some have gone to biometric systems that take attendance automatically using things like iris scans, fingerprints, or face recognition, while others are stuck on paper records or files. In the realm of image processing, face recognition ranks high among the most common uses. Because of this, its popularity has grown exponentially over the last few years. Haar cascade, histograms of local binary patterns, Eigen faces, and Fisher faces are among the face recognition and identification techniques that are part of the Open CV package. Part one of the process entails locating, normalising, face detecting, and aligning; part two is extracting and matching features; and part three is recognition.

II. RELATED METHODS

The research [1] identified the viola jones detector as the most successful approach for multiview and frontal face recognition. The detector employed the Haar features to properly extract the most crucial traits, such as the eyes, mouth, nose, and ears. The suggested method outperforms rival algorithms in terms of resilience and face orientation identification, even though the training sample was quite small. Nevertheless, these methods could be time-consuming to train and don't do well in environments with varying illumination. The study's automated attendance system was created by merging the LBPH (Local Binary Pattern Histogram) method with the Haar cascade technique [2]. The tracked image is compared to the trained picture after Haar Cascade image training. After using the LBPH algorithm for facial recognition, the subsequent step is to take attendance. Because it consistently yields accurate results in all lighting conditions, the Haar Cascade approach is both efficient and reliable. An accuracy rate of around 96.68% is achieved by the procedure. While the study did examine the accuracy of various approaches, the LBPH strategy proved to be the most successful, which is a drawback of the research. A well-lit classroom may help students focus and complete assignments more quickly and accurately. Facial characteristics may be extracted using facial vectors by combining the Eigen face methodology [3] with the SVM method. In this work, we train the dataset to perform detection and classification. Support vector machine, which was also provided, was also used by the author. Despite the technique's efficiency increasing when the dataset includes more faces, it does have a downside: using the SVM method, it only achieves 61% accuracy with 50 real-time images. For the method to be refined, the dataset has to include more images of faces. With a 98.3% success rate, the CNN approach is much superior than its rivals. The proposed system in this study [4] recognises faces and assesses their effectiveness using methods such as SVM, Haar classifiers, KNN, and CNN. Reports will be generated and stored in the database after face recognition. Lighting, head movement, and changes in pupil-to-camera distance are just a few of the many variables tested to determine the system's efficacy. The KNN approach outperformed two other algorithms that were evaluated for the system's advantages due to its minimal installation need, 99.27% accuracy rate, and cost-effectiveness. In the research, Convolutional Neural Networks (CNNs) [5] were used for face detection, landmark identification, and recognition. Nevertheless, there are a few of drawbacks to using CNNs, including the SVM method's inefficiency and the low processing complexity. It is crucial to provide many images of a person's face so the system can recognise them before adding their name to the Excel sheet. When given an image as input, the system's primary goal is to determine where the image came from.

The approach may have limitations, such as the fact that it can only identify certain faces in photos. To get the metrics

International Journal of Advanced Trends in Multidisciplinary Research
ISSN: 3139-0986, Volume 2 Issue 7
July 2026, PP 1-5

for the degree of similarity between two faces, you must first locate each one. After then, you'll be able to compare them. The fact that we built the system using Node.js, Face recognition.js, and dlib is a beneficial part of the research. Turn the camera using the dlib library (open CV) to focus on certain features or the whole face. Both the quickness and quality of its identification are exceptional. Here, we demonstrate how to implement ensemble learning for face recognition [6]. The approach used in this research was ensemble learning using convolutional neural networks. All three of the model's subnet networks—AlexNet, VGG Net, and ResNet—were used by an SVM classifier to formulate the final forecast. The method and model have several advantages, such as a short prediction timescale, good facial recognition capabilities, and a high identification accuracy of 71.27%. Some of the limitations of the model, such as its poor face recognition rate and 60% accuracy rating, are highlighted in the research when compared to other approaches like transfer learning. With regard to the recognition feature vector, this study examines how well the t-distributed stochastic Neighbour Embedding method performs when it comes to gender categorisation [7]. The identifying feature vector is classified in this work using a variety of machine learning approaches, including K-closest neighbour, Support vector machine, Random Forest, and Logistic Regression, among others. Across a wide variety of datasets, our proposed method has consistently achieved a high identification rate. Specifically, 99.2%, 98.7%, and 97.4% were the equivalent percentages for the face datasets. While good features may directly raise task recognition rates, the vast majority of invalid features would still have a noticeable influence on recognition effects in this case, which is a drawback of feature extraction studies. Two face recognition algorithms, the Local Binary Pattern method and the Haar cascade technique, are compared in this paper [8]. The main objective of the system is to find the total number of faces in the picture. The Haar cascade is superior than the LBP method in terms of accuracy and the amount of faces it can detect in an image, which is a positive aspect of the procedure [9]. Although it isn't nearly as precise, the LBP classifier's technique is quicker than the Haar cascade classifier's. To achieve face recognition, the author of the research [10] used K-nearest neighbours' techniques, Successive Mean Quantisation Transform methodology, and deep recurrent learning. Researchers focusing on live video facial recognition may find this material valuable. Though it struggles with low-quality photos, the model's ability to isolate user faces from large groups of people is a strength. Face identification is useless in images that include distortion, blurriness, raggedness, or visibly damaged hinged parts. In their analysis of three algorithms for mask detection, the authors of the aforementioned research [10] conclude that KNN, SVM, and Mobile Net provide the best results. An advantage of the method is that it maintains a high accuracy rate of 88.7 percent by using the best algorithm. When compared to SVMs and KNNs, it is more effective in decreasing the network size and parameter count. Furthermore, it is quick and works well with mobile apps. One limitation of the research is that our model had a higher accuracy rate (78.1% vs. KNN and SVM). The face detection section of this study uses Face Net for face recognition and is based on the MTCNN approach. In addition, we can identify whether the user blinks by using an Ensemble of Regression Trees for face liveness detection. We propose this approach to identify when a user is present and when their face is active. The results of the trials demonstrate the technology's high level of accuracy, which stands at 99.9 percent. One issue with the system is roll. More than a 20-degree tilt of the head will prevent it from recognising you. Using face recognition software to keep tabs on pupils' whereabouts is borne up by the results of this research [12–16]. The findings show that SVC is great at classifying faces while Faster R-CNN is great at discovering new faces. In addition, face embedding has been created using Face Net. Among the many advantages of RPN for ROI extraction is its high rate of face recognition, which achieves an overall accuracy of 93.35 percent. The method's inability to rapidly identify the whole scene using RPN is a drawback. The suggested system uses multi-trained convolutional neural networks (MTCNNs) to automatically identify and categorise students' faces, allowing for the tracking of their attendance. What makes the suggested approach important • The suggested approach can recognise a large number of faces. The three-stage MTCNN approach is used to accomplish face classification. • The likelihood of human mistake is eliminated when student attendance is automatically tracked. The process of the suggested approach is shown in Figure 1.

III. PROPOSED METHOD

The three steps that make up this proposed method are: 1) finding one or more faces inside an image; 2) picking out important features from those faces; and 3) recognising a face in an image by studying a subset of human facial characteristics such the eyes, nose, and mouth. We shall adhere to a multi-stage process to accomplish this. Look at Face's Versatility We will use the embedding model that Cascaded Convolutional provides to locate the facial area. It does a basic alignment on every face it finds. The suggested method's stages are shown in Figure 2.b. Phase one of data collection is gathering all pertinent details: Pull the faces away from the camera by pressing. Our technology watches the user while they use the GUI and uses that information to identify their face.

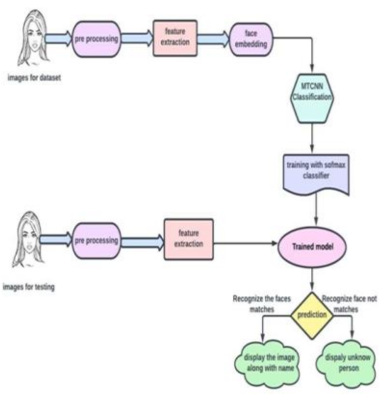

Fig 1. Proposed Method

International Journal of Advanced Trends in Multidisciplinary Research
ISSN: 3139-0986, Volume 2 Issue 7
July 2026, PP 1-5

I went ahead and captured 500 photos so I could train the dataset and test up to 10 people. All of the faces that have been gathered are immediately stored in a folder with their names. An ANN with Convolutional Layers for Multiple Use Cases (B) The study recommends using multi-task learning, or MTCNN, to combine the aligning and recognising tasks. At first, it quickly uses a shallow CNN to produce candidate windows. Using a more advanced CNN to improve the suggested candidate windows is the second step. As a last step, the third stage employs a more sophisticated CNN to refine the output and localise face features. An method known as a Multitask Convolution Neural Network (MTCNN) [13] may be used to recognise faces and other facial features in pictures. Facial detection may be achieved using MTCNN. You may trust it and rely on its accuracy. Face identification maintains accuracy in the presence of lightning, huge size shifts, or extreme rotations. Locating facial landmarks, performing bounding box regression, and classifying faces as opposed to non-faces are the three main characteristics of Markov TRNN. The five most noticeable characteristics of MTCNN are the nose in the face, side eyes (both left and right), corners of the mouth (both left and right), and nose in the face. Combine OpenCV with MTCNN to create a facial recognition system. Steps to Implement an MTCNN Then, an image pyramid, which is the result of scaling the picture, is fed into a three-stage transmission network. A thick layer is an integral part of the design of the Proposal Network (P-Net). The bounding box regression vectors of the candidate windows may be obtained using this proposal network. Bounding box regression is a popular approach to facial object recognition; it uses a previously defined class to predict where boxes will be placed. Acquiring the bounding box vectors allows for the refinement and subsequent joining of the overlapping areas. After the total number of applications has been reduced and polished, this step delivers all candidate windows as its final output. Second Step: R-Net, or the Refinement Network The Refine Network receives all of the data supplied by the P-Net.

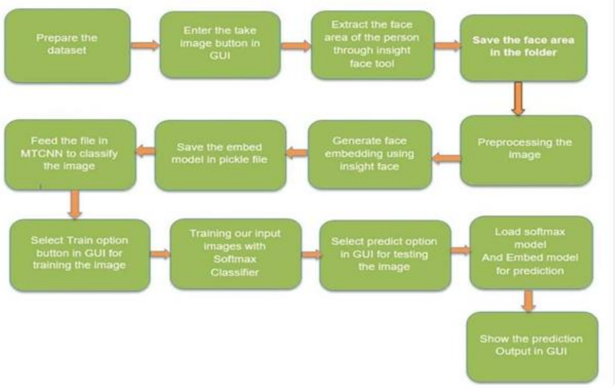

Fig2. Proposed method work flow

The reason for this is because the network design ultimately produces a thick layer. Bounding box regression is used by R-Net for calibration. After that, it further narrows the pool of choices by integrating overlapping ones using non-maximum suppression (NMS). In the case when the input is a face, the R-Net produces a vector with ten elements for localising facial landmarks, and another vector with four elements for the face's bounding box. Finally, the Output Network, or O-Net By pinpointing the exact locations of the five most important facial features—the mouth, nose, eyes, and eyebrows—the Results Network aims to provide a more comprehensive picture of the face. At this stage, similarities with the R-Net become apparent.

Face Embedding Insight

There are several uses for face, a sophisticated facial analysis tool, including face alignment, detection, and recognition. The insight face model may be used to create picture embeddings. A 128-value matrix extracted from a face shot serves as input to the Insight Face model, which prioritises certain facial characteristics. This vector is called an embedding in the machine learning domain. All of a picture's crucial elements are included in this vector.

Training

With the help of the GUI, we can start training the picture. The next step is to use the face embedding to categorise the individual. This is done by feeding the input picture into an input layer, which then teaches the hidden layer the embedding feature. Lastly, the learnt feature is sent to an output layer that has the SoftMax function. This will help us categorise the individual. Try out F: The camera prediction feature of the facial recognition system is accessible via the user interface. Both the SoftMax and Embed Pickle models will be loaded. The face embedded areas will be used by the Embed model to find the face area in the camera. The SoftMax model will try to identify the face as a student after it discovers the region of the face, and the system will then record the data of the related student.

IV. IMPLEMENTAION & TESTING

Once we launch the camera on our own computer, our model will use a tool to eliminate the face area. Subsequently, video processing will begin. A folder tagged with the subjects' names will automatically be created to store the gathered face photos. Here is how the dataset was collected. Each user's folder must have no less than 500 photographs according to our method. People with spectacles or beards are just two of the many possible topics it might capture images of. For training, we have gathered ten people's images, and for testing, we have F-tested thirty percent of those pictures [14][15][16]. Several measures are used to evaluate the model's performance and quality. Performance metrics or assessment metrics are the names given to these measures. This metric demonstrates the efficacy of our approach when applied to the provided data. It is possible to improve the model's performance by making these changes to the hyper parameters.

The success of DL models hinges on how well they deal with fresh data. Performance metrics may be used to assess how well a model performs on the updated data. See how the proposed solution compares to others and the metrics it employs in Table 1. By dividing the total number of

International Journal of Advanced Trends in Multidisciplinary Research
ISSN: 3139-0986, Volume 2 Issue 7
July 2026, PP 1-5

predictions by the total number of samples, we can get the prediction accuracy (Acc) from equation (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

The model's sensitivity (S), defined as the proportion of true positives it correctly recognises, is given by Equation (2).

$$Sensitivity\ (s) = \frac{TP}{TP+FN} \quad (2)$$

Precision (TPR) is the percentage of positive outcomes that the model accurately recognised as positive, as shown in equation (3).

$$Specificity = \frac{TN}{TP+TN} \quad (3)$$

The number of relevant occurrences retrieved, denoted as precision P, is obtained from Equation 4.

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

A comparison of the Haar cascade CNN algorithms is shown in Table 1.

Table 1: Comparision Of Mtcnn With Haar Cascaded With Cnn

	MTCNN	HAAR CASCADE WITH CNN CLASIFICATION
ACCURACY	98.2%	84%
SENSITIVITY	99%	83.7
SPECIFICITY	97.5%	99%
Precision	97.5%	95#

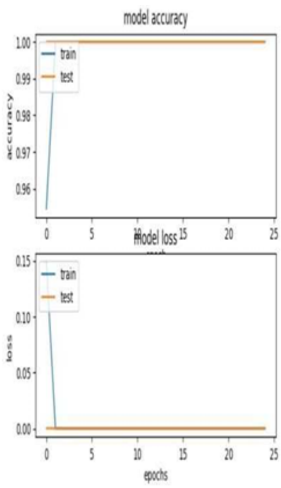

Fig 3. Accuracy of the proposed method

The accuracy of the proposed method is shown in Figure 3, and a typical outcome from the proposed method is shown in Figure 4.

Fig 4.Result of the proposed System

Table 2: Comparision with Other Methods

METHOD	ACCURACY
LOCAL BINARY PATTERN (LBP)[13]	95%
DUAL VARIATIONAL GENERATION FRAMEWORK [15]	97.21%
OPENFACE FACE RECOGNITION LIBRARY [14]	85.70%
MTCNN (PROPOSED METHOD)	98.2%

In Table 2 we can see how the proposed method compares to others. Table 2 displays the better outcomes produced by the MTCNN. The three parts of MTCNN work together to create a cascaded architecture that improves the accuracy of face recognition as it goes along. Among its many applications, the MTCNN architecture IV produces bounding boxes.

V. CONCLUSION

This is one way that teachers and TAs can keep tabs on who showed up to class, lab, or session. It will save a tonne of time, especially when there are a lot of students in the class. A suggested automatic attendance system aims to fix the problems with the current method. Using this system, a teacher or TA may see how many people are present in a class, lab, or other meeting. It will save a tonne of time, especially when there are a lot of students in the class. A suggested automatic attendance system aims to fix the problems with the current method. This is a method for tracking students' presence in class that uses picture analysis. Aside from facilitating attendance monitoring, this technique also holds the promise of enhancing the school's reputation. The suggested MTCNN performed better than competing systems. The student will be notified via text message to remember to bring their attendance card in the future. Using a GSM module makes it possible. While a kid is in class, a parent can get a text message.

REFERENCE

1. CY. Chen, JJ. Ding, and HW. Hsu. “Prominent facial feature and hybrid learning method-based advanced face detector robust to head-up, head down, and arbitrary rotation cases”. Signal, Image and Video Processing, Vol. 15, pp. 147-54. 2020.
2. S.S. Pawaskar, and A.M. Chavan. “Face recognition-based class management and attendance system”. Proc. IEEE Int. Conf. Bombay Section Signature Conference (IBSSC), pp. 180-185, 2020.
3. M. Geetha, R. S. Latha, R. S, S. K. Nivetha, S. Hariprasath, S. Gowtham, C. S. Deepak. “Design of face detection and recognition system to monitor students during online examinations using Machine Learning algorithms”. In 2021 international conference on computer communication and informatics (ICCCI) IEEE, pp. 1-4. Jan. 2021.
4. S. Dev and T. Patnaik.” Student attendance system using face recognition”. International conference on smart electronics and communication (ICOSEC) IEEE, pp. 90-95. 2020
5. H. Klym, and I. Vasylchyshyn. “Face detection using an implementation running in a web browser”. IEEE 21st International Conference on Computational Problems of Electrical Engineering (CPEE), pp. 1-4, 2020.
6. C. Jia, C.L.Li, and Z. Ying, “Facial expression recognition based on the ensemble learning of CNNs”, IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), pp. 1-5, Aug 2020.
7. Y. Lin, and H. Xie. Face gender recognition based on face recognition feature vectors. IEEE 3rd International conference on information systems and computer aided education (ICISCAE), pp. 162-166, Sep 2020.
8. A.B. Shetty and J. Rebeiro. “Facial recognition using Haar cascade and LBP classifiers”. Global Transitions Proceedings, Vol 2(2), pp. 330-335. 2021.
9. H. L. DhanushGowda, K. Vishal, B.R. Keertiraj, NehaKumariDubey, M. R Pooja.” Face Recognition based Attendance System”, International Journal of Engineering Research & Technology (IJERT), Vol. 9 (06), June 2020
10. A, B, Shetty, and J. Rebeiro. “Facial recognition using Haar cascade and LBP classifiers”. Global Transitions Proceedings, vol. 2(2), pp. 330-335. 2021
11. A. Singh, “Face Detection using Deep Recurrent Learning and SMQT Technique”. International Conference on Electronics and Sustainable Communication Systems (ICESC). pp. 17- 22. 2020
12. W. Vijitkunsawat, and P. Chantngarm, “Study of the performance of machine learning algorithms for face mask detection”. International conference on information technology (InCIT), pp. 39-43, August 2020.
13. S. Huang, and H. Luo, “Attendance system based on dynamic face recognition”. International Conference on Communications, Information System and Computer Engineering (CISCE). pp. 368-371 July 2020.
14. B. Amos, B Ludwiczuk, and M. Satyanarayanan, M. “Openface: A generalpurpose face recognition library with mobile applications”. CMU School of Computer Science, Vol 6(2), 2020.
15. S.M. Bah, and F, Ming, “An improved face recognition algorithm and its application in attendance management system”. Array, 5,2020.
16. C. Fu, X. Wu, Y, Hu, H. Huang and R. He . ” DVG-face: Dual variational generation for heterogeneous face recognition.” IEEE transactions on pattern analysis and machine intelligence, 44(6), 2938- 2952, 2020.