Real time Face Detection and Optimal Face Mapping for Online Classes

Archana MCP¹, Nitish CK², Sandhya Harikumar³
¹ ² ³ Department of Computer Science Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India
E-mail: ¹ anair11295@gmail.com
E-mail: ² nitishck1994@gmail.com
E-mail: ³ sandhyaharikumar@am.amrita.edu

Abstract. The main objective of this paper is to provide a web-based tool for identifying faces in a real-time environment, such as Online Classes. Face recognition in real-time is now a fascinating field with an ever-increasing challenge such as light variations, occlusion, variation in facial expressions, etc. During the current pandemic scenario of COVID-19, the demand for online classrooms has rapidly increased. This has escalated the need for a real-time, economic, simple, and convenient way to track the attendance of the students in a live classroom. This paper addresses the aforementioned issue by proposing a real-time online attendance system. Two alternative face recognition algorithms are perceived in order to develop the tool for real-time face detection and recognition with improved accuracy. The algorithms adopted are Local Binary Pattern Histogram (LBPH) and Convolutional Neural Network (CNN) for face recognition as well as Haar cascade classifier with boosting for face detection. Experimental results show that CNN with an accuracy of 95% is better in this context than LBPH that yields an accuracy of 78%.

Keywords: Face Detection, Face Recognition, OpenCV, LBPH, Haarcascade, CNN.

1. Introduction
Surveillance systems have become inevitable for securing various places including home, hospitals, banks, and airports. Such systems require installation of surveillance cameras for constant monitoring of the activities of people. One of the primary functionalities of these systems is face recognition in real-time. With the surge of COVID, the need for online classes and online proctoring in education systems has been adopted with various challenges. One of the biggest challenges that teachers face is real-time attendance of the students while the live classes are being held. Biometrics, based on behavioural and physical characteristics, are employed in the majority of corporate and government authentication systems¹-⁴. The biometric system converts raw data such as a person’s face, fingerprint, or iris into valuable features. A feature describes the sort of data provided to the biometric system⁵ and how it is used to make decisions. Because of the variability present in human faces, such as pose, expression, position and orientation, color of skin, the presence of facial features like hair, differences in camera specifications, variation in light, and image resolution, the method of face detection in real time is complicated. The technique of detecting an object in an image is known as object detection.
For example, in face detection methods, a human face is detected from an image including various objects such as cars, buildings, and so on. Face Recognition can be done using supervised and unsupervised technique. The first and the essential step for any face recognition system is face detection. Face detection is done using a face detection algorithm to identify and extract faces from an image. Face detection methods include knowledge-based, feature-based, template matching, and appearance-based techniques. A face-based rule is described by a knowledge-based or rule-based technique. It decreases computing complexity, but the rate of face detection is lower as only frontal faces can be identified, and rule creation is complicated. The detection is difficult due to the complex background and many faces. These problems are handled by feature-based methods, which employ facial characteristics such as the nose, eyes, mouth etc. to detect features. This approach is quick, reliable, has a high detection rate, and has a low computational cost and complexity. It is first trained as classifier then used to differentiate positive and negative images. 94% success rate is obtained using this method. But, in some cases convolution cost is high. The disadvantages of the feature-based technique are overcome using the appearance-based technique. The characteristics of face images are understood using statistical analysis and machine learning in this technique. This approach may be used to extract features for face recognition. Face detection, emotion detection, face recognition, object detection, and other detection issues are effectively implemented by using appearance-based techniques.

In this work, for a real-time attendance system, LBPH, a feature-based technique, and CNN, an appearance-based technique, are employed. Due to the prevalence of online classrooms, it is increasingly difficult to determine whether or not students are present in the classroom. This work can be utilised as a user-friendly attendance system to tackle this problem. The attendance will be taken only if the student’s face is visible during the class. This allows the instructor to be notified of a specific student’s absence via an alert. The main contributions are:

- Developing efficient real time face recognition with minimal error rate by adapting two algorithms namely LBPH and CNN
- An interactive web application for real time online attendance system.

2. Literature Survey
Various algorithms for face detection and recognition are available. The study by Viola and Jones offers a machine learning strategy for visual object recognition that can analyse pictures quickly and achieve high detection rates. There are three primary contributions in this study. The first contribution is to use integral pictures rather than regular photos. The second contribution is to use the AdaBoost method to combine weak classifiers to generate a strong classifier. The use of cascade to identify the perfect object and acquire the region of interest is the third contribution. Detection is improved with this technique, regardless of skin tone. Face identification and preprocessing using the eigenvector and eigenface concepts are proposed in, with the result being fed into ANN and PCA algorithms. Finally, the algorithms are compared to kmeans and Fuzzy c-means. However, it is ineffective for side postures, and visualisation is tough. Our work is inspired by the ideas proposed in 10 and 11 to employ CNN and Local Binary Pattern Histogram (LBPH) respectively for facial recognition. LBPH algorithm can distinguish both the front and side faces. The pixel value in the centre was modified with the median of the neighbourhood value, and the feature was retrieved using sub blocks and histograms. Deep learning methods extract different facial features. The approach is mostly centered on CNN. In two layers, there are two normalizing operations. Batch normalisation is the method used to speed up the network. Finally, the picture is predicted using the Softmax classifier. Batch normalisation has a drawback of computational overhead during training. Face
Recognition Based Attendance System Using Machine Learning Algorithms such as SVM, MLP, and CNN, to detect the faces in videos and photos is proposed in [13]. However, SVM method is ineffective on huge datasets as it leads to underfitting if the number of features for each data point is greater than the number of training data samples. CNN led to more accurate recognition as compared to other algorithms. The authors in [14] proposed the real-time attendance system where face enrollment is done and stored in the database. Face recognition is performed here utilising a feature-based skin classification approach. However, this is not a good approach for recognising faces because factors like face sizes, poses, lighting conditions affects the performance. The proposed technique in this research aims to construct a face detection and identification system that is not affected by light intensity, facial position, occlusion, or low resolution. This technique developed a high-accuracy, high-speed real-time face detection tool with an interactive user interface. An effective and user-friendly online attendance web application is also built to ensure flawless online attendance in the classroom. Unlike other existing methods, here feature based and appearance based techniques are used.

3. Proposed Work
There are primarily two approaches for recognising faces LBPH and CNN. Fig. 1 depicts a high-level design of the proposed approach.

![Figure 1. Work Flow of Real-time Online based Attendance System.](image)

3.1. Haar-cascade classifier
The Haar-cascade classifier is used for face detection using OpenCV. In this system, the OpenCV library is utilised to process images for real-time face identification. The Haar-cascade is a machine learning technique that involves training a cascade function using a large number of positive and negative images. Those with faces are considered positive, whereas images without faces are considered negative. Image characteristics are viewed as numerical information taken from images that can distinguish one image from another in face detection.
The first stage in this is to collect Haar-like features. The detection window is used to calculate a Haar-like feature. A rectangle zone will appear in the detection window at a specific location, and required features will be taken from this area. The sum of the intensity from each region will be added, and the difference between the sums will be found. However, for larger pictures, this is not possible, thus we use integral images. Integral pictures aid in the discovery of Haar like features. It constructs sub-rectangles and array references for each rectangle, but it does not use the pixel intensity evaluation method. The Haar like features are then computed using these. We need the best feature out of these features. For this we utilise the Adaboost algorithm.

Adaboost is one of the algorithms used to detect faces. It generates iterative ensembles by integrating multiple weak classifiers to improve system accuracy and obtain a strong model for classification. The main idea behind the Adaboost classifier is to set the weights of the classifier and train the data points iteratively to ensure accurate predictions of uncommon observations. Adaboost uses the best features to train a classifier. To detect the face, a weak classifier is utilized to construct a strong classifier. Weak classifiers are built by moving the window around in the input image. It determines the proper positive or negative threshold for categorising the faces. A large number of Haar-like features are required to develop a strong classifier because these are weak classifiers.

The Adaboost algorithm is represented as in Equation 1:

\[ A(x) = \alpha_1 a_1(x) + \alpha_2 a_2(x) + \alpha_3 a_3(x) + \alpha_4 a_4(x) + ... \]  

\( a_i(x) \) is the class point of \( x \) assigned by \( i^{th} \) weak classifier

\( A(x) \) is the class point of \( x \) assigned by strong classifier

The final stage is to utilise a cascade classifier to combine weak learners to create strong learners. The cascade classifier is depicted in the diagram below.

![Cascade Classifier Diagram](image-url)

**Figure 2.** Design of Cascade Classifier.

There are weak learners in each level of the cascade classifier. To train these learners, adaboosting is utilised, which results in a high-accuracy strong classifier that uses the mean of different weak learners. The importance of phases is to quickly retrieve negative pictures since the majority of the image will be the area without a region of interest. Misclassifications and errors do occur. We select attributes with the lowest error rate, implying that they are the most effective at distinguishing between facial and non-facial images. Face recognition is done using the LBPH and CNN algorithm.
3.2. Local Binary Pattern Histogram

LBPH algorithm is a straightforward method for labelling image pixels and thresholding their surroundings. To put it another way, LBPH summarises an image’s local structure by comparing each pixel to its neighbours and converting the result to a binary number. The goal of this method is to extract local information from photos. The primary principle is to focus on the local features of an object rather than the entire image as a high-dimensional vector. In every illumination environment, the LBPH algorithm is applied.

![Figure 3. Steps of real time Facial recognition.](image)

Workflow of the LBPH algorithm is shown in Fig. 3. Here, a 3x3 window is evaluated, and it is moved over the entire picture. As the window advances, the pixel in the centre is compared to the pixel in the surrounding area. The threshold is determined by the intensity value of the centre pixel, and the surrounding pixels are represented in binary format with regard to the threshold. The value is expressed as 1 if it is greater than or equal to the threshold, and as 0 if it is less than the threshold as shown in the Fig. 4.

When the binary values are read in a clockwise direction, a binary pattern like 00111110 emerges. This technique is repeated across the image, resulting in a large number of distinct binary values. These binary numbers are converted to decimal values, and a histogram for that image is obtained as shown in Fig. 5.

Create a model using the same procedure for the entire dataset. If a new picture is put into the model, it will generate an LBPH and compare the generated histogram to the trained model’s histograms. Finally, it identifies the best match and assigns a label as well as a confidence level.

3.3. Convolutional Neural Network

Multiple layers of artificial neurons make up convolutional neural networks. A Convolutional Neural Network is a Deep Learning method that can take an input image and assign relevance
weights and biases to distinct items in the image, allowing it to distinguish between them. When compared to other classification algorithms, the amount of pre-processing required by a CNN is significantly less, has the benefit of requiring less memory and less parameters. As a result, the performance is excellent.\textsuperscript{12}

The proposed CNN architecture for face recognition has five convolutional layers, as shown in the Fig. 6.

\textbf{Figure 5.} Steps of LBPH Algorithm

\textbf{Figure 6.} Proposed architecture for a face recognition system using CNN.

The system’s output will be less in size. The features will be presented in a linear fashion. The ReLU (Rectified Linear Unit) activation function is used on attributes that increase non-linearity. The system will be unstable due to linear properties. ReLU is calculated using Equation 2:

\begin{equation}
X_r = ReLU(A_k) = \max(0, A_k)
\end{equation}

\((A_k)\) is the pixel at \(k^{th}\) location

The pooling layer is used to preserve the image’s characteristics. There are other sorts of pooling layers, but max pooling is applied here. We can lower the size of the representation and boost the speed of computation by implementing pooling layers. The flattening method is
used to decrease an image’s dimensionality, such as converting a 3-D image to a 2-D plane. This information is subsequently forwarded to the fully connected layer.

The fully connected stage of CNN design connects each and every neuron in one layer to a neuron in another layer. It’s really similar to MLP (Multilayer Perceptron). The fully connected layer provides the final forecast output. Softmax is the activation function in this case. Softmax will return a binary value of 0’s and 1’s as a result. Probability of softmax is as follows:

\[
P(A_i) = \frac{\exp(B_i)}{\sum_{j=1}^{n} \exp(B_j)}
\]  

(3)

3.4. Django Framework based User Interface

Django is a Python-based web framework. It is popular among the more experienced web developer community. Django Framework architecture is based entirely on Model, View and Controller (MVC), which is designed for “perfectionists with deadlines”. Because of the well-structured architecture of Django web application development, it is very quick and well supportive in agile web application development methodology. The Django template engine is well-structured. We can use the Django template engine to reduce or eliminate the need for frontend frameworks such as AngularJS, Vue.js and others.

Figure 7. GUI of Face Recognition system.

User interface of proposed system is depicted in Fig. 7. We choose the type of classifier, then click the Detect Face button. A model with live video streaming will open, and the person will be identified based on the trained model.

4. Result and Analysis

4.1. Experiment Setup

In this work, the model is trained using images of each individual. We trained our model on two data sets. The first dataset is built with the use of a laptop camera. Videos of people facing in distinct directions are captured. Face detection and frame extraction are then performed on these videos. There are 1000 photos in each class. Each image is downsized to 50x50 pixels to keep up the compute limitations. The second is ORL public dataset. The suggested design system was run on an Intel(R) Core(TM) i5-1035G1 quad-core CPU, 8.00GB of RAM, and a 1 MP camera to capture the image. Pycharm is used as the development environment for creating a web application based on the Django framework. It contains interactive user interface. The CNN model is trained and created using Anaconda Navigator. One of the applications used for online classrooms is Microsoft Teams. The proposed web application can be embedded in Microsoft Teams as a channel so that it is used for attendance marking system.

In this section results and the impact is illustrated.

- Recognition accuracy improves as the amount of training data increases as shown in Fig. 8
When the size of the training images varies, the model’s training accuracy decreases. As a result, it has an impact on the system’s performance. To solve this issue, we should resize the images to a specific size (50x50).

- It has been discovered that when CNN contains 4 hidden layers, the system is more accurate as shown in Fig. 9.
- The amount of time taken for training is proportional to the number of hidden layers as shown in Fig. 10.

**Figure 8.** Graph depicts Training data v/s Accuracy plot. Accuracy is directly proportional to data.

**Figure 9.** Graph shows Training Layer v/s Accuracy plot. In the fourth hidden layer accuracy is high.

**Figure 10.** Graph shows Training Layer v/s Time plot. Training layer and training time is directly proportional.

### 4.2. Face Recognition

Fig. 11 shows the results of the Convolutional Neural Network model prediction. Here, 20 random images are recorded, and the images are then tested using the model that was constructed, which has five hidden layers. It was also discovered that 95% of the images were appropriately identified and presented using matplotlib. Recognized name is also mapped in the image.

For testing the accuracy of the system we used ORL public dataset. There are 400 photos in the ORL Database of faces from 40 different people. The photographs for certain persons were shot at several times, with varied lighting, face emotions (opened eyes, smiling), and face features being used. The people were photographed in an upright, frontal stance against a black, uniform background (with tolerance for some side movement). Each picture is 92x112 pixels in size and has 256 grey levels per pixel. Result of trained model recognition using ORL is shown in Fig. 12.
Figure 11. Correctly recognised faces using 20 random faces from the test data and acquired 95% accuracy.

Figure 12. Correctly recognised faces from ORL database having 40 different people.

4.3. Real time face recognition using Django web Application

Figure 13. Real time Face Recognition system using LBPH algorithm

Figure 14. Real time Face Recognition system using Convolutional Neural Network.

5. Conclusion
Face recognition is a difficult yet significant recognition technique that has one major benefit of being user-friendly over all other biometric techniques. In this paper, we present a web-based tool for real-time face recognition utilising CNN and LBPH. We found that in CNN, the accuracy is 95%, while in LBPH, the accuracy is 78%. Proposed method is economic and can be deployed on commodity laptops.
Future scope lies in the usage of high-quality webcam to increase the reliability of face recognition system. Further, faces of a larger number of people in diverse contexts can also be utilised to improve identification accuracy. NoSQL/SQL databases can be used to improve training performance of the model. This technology can be utilised as a real-time attendance system as well as a facial expression analysis system to assist teachers in properly managing online classes.

References

[1] Vivek Menon, Bharat Jayaraman, and Venu Govindaraju. Enhancing biometric recognition with spatio-temporal reasoning in smart environments. *Personal and ubiquitous computing*, 17(5):987–998, 2013.

[2] MS Sumi Suresh, Athi Narayanan, and Vivek Menon. Maximizing camera coverage in multicamera surveillance networks. *IEEE Sensors Journal*, 20(17):10170–10178, 2020.

[3] Sandhya Harikumar, Viveka Mannam, Chiranjeeb Mahanta, Mounika Smitha, and Shazia Zaman. Interactive map using data visualization and machine learning. In *2020 6th IEEE Congress on Information Science and Technology (CIST)*, pages 104–109, 2020.

[4] Nilesh A Lal, Salendra Prasad, and Mohammed Farik. A review of authentication methods. *vol*, 5:246–249, 2016.

[5] Vivek Menon, Bharat Jayaraman, and Venu Govindaraju. Integrating recognition and reasoning in smart environments. In *2008 IET 4th International Conference on Intelligent Environments*, pages 1–8. IET, 2008.

[6] Sandhya Harikumar and Raji Ramachandran. Hybridized fragmentation of very large databases using clustering. In *2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)*, pages 1–5. IEEE, 2015.

[7] P Viola and M Jones. Face detection. *IJCV*, 57:2, 2004.

[8] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, volume 1, pages I–I. IEEE, 2001.

[9] Mayank Agarwal, Himanshu Agrawal, Nikunj Jain, and Manish Kumar. Face recognition using principle component analysis, eigenface and neural network. In *2010 International conference on signal acquisition and processing*, pages 310–314. IEEE, 2010.

[10] Reny Jose. A convolutional neural network (cnn) approach to detect face using tensorflow and keras. *International Journal of Emerging Technologies and Innovative Research, ISSN*, pages 2349–5162, 2019.

[11] XueMei Zhao and ChengBing Wei. A real-time face recognition system based on the improved lbph algorithm. In *2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP)*, pages 72–76. IEEE, 2017.

[12] Musab Co¸skun, A. U¸car, Özal Yildirim, and Y. Demir. Face recognition based on convolutional neural network. *2017 International Conference on Modern Electrical and
[13] Radhika C Damale and Bazeshree V Pathak. Face recognition based attendance system using machine learning algorithms. In 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), pages 414–419. IEEE, 2018.

[14] K Senthamil Selvi, P Chitrakala, and A Antony Jenitha. Face recognition based attendance marking system. International Journal of Computer Science and Mobile Computing, 3(02):337–342, 2014.

[15] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. Understanding of a convolutional neural network. In 2017 International Conference on Engineering and Technology (ICET), pages 1–6. Ieee, 2017.

[16] Ralph Gross. Face databases. In Handbook of face recognition, pages 301–327. Springer, 2005.