

Dense Feature Based Face Recognition from Surveillance Video using Convolutional Neural Network

K. Suma^a, N.V. Sunitha^b, N. Suhasini^c, T. Shreekumar^d

^aFaculty in Computer Science, Mangalore Institute of Technology and Engineering, Moodabidri, India.
E-mail: sumaraghuraj@gmail.com

^bFaculty in Computer Science, Mangalore Institute of Technology and Engineering, Moodabidri, India.
E-mail: sunithanv6720@gmail.com

^cFaculty in Computer Science, Mangalore Institute of Technology and Engineering, Moodabidri, India.
E-mail: suhasini.nellingery@gmail.com

^dFaculty in Computer Science, Mangalore Institute of Technology and Engineering, Moodabidri, India.
E-mail: shreekumart@gmail.com

Article History: Received: 11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2021

Abstract: Face Recognition is a field of identifying the person from the facial features and has wide application range in security, human computer interactions, finance etc. In recent years, many researchers have developed different algorithms to identify the Faces from various illumination variations and Pose variation, but these two problems remain unsolved in Face Recognition (FR) field.

The Local Binary Pattern (LBP) has already proved its robustness in illumination variation. This paper proposes a four-patch Local Binary Pattern based FR utilizing Convolutional Neural Network (CNN) for identifying the Facial images from various illumination conditions and Pose variation.

Keywords: Face Recognition, Four-patch Local Binary Pattern, Convolutional Neural Network.

1. Introduction

In the field of biometric research, face detection and Face Recognition are the main contents and in the field of pattern recognition it is a frontier subject. Face Recognition is an important method used to identify an individual which is largely used in security, health care system, criminal identification, surveillance operations, personal identification, document verification and many other fields. Face Recognition is more convenient and has the validness compared with other methods like fingerprint, palm, retina and iris. In the Face Recognition method person need not cooperate actively. The accuracy of the Face Recognition technology depends on the conditions such as frontal faces and indoor lighting conditions [1]. For the uncontrolled conditions (outdoor light, different face angles) the Face Recognition task is still not accurate up to the requirement. Proper technique is required for face detection and Face Recognition to handle different expressions of the face, variations in pose, factors of aging and resolution either in the frame of stationary objects or video images. In Face Recognition technique resolution plays a major role in identifying a face in surveillance or CCTV. In Face Recognition system, first we have to identify a face in an image. The major subject of face detection is to identify whether there is a face in an image or not, in a view of stationary image or video image.

In most of the automatic vision systems, detection of object is the first step. Face detection algorithms are developed and trained to robust detection and accurately locate faces or objects within the images[2]. These can be in real time from a video camera or from photographs. To recognize face, software must first detect the face and identify the features before making an identification. Current researches are focusing on how to improve the performance of detection in unrestricted conditions. In our work using Haar Cascade Classifiers the face is detected.

Face Recognition is a very important application of computer vision and popular biometric technique mainly used in security which basically validate the identity of the user. Face Recognition technology is one of the fastest growing technology in biometric field. In this method most unique characteristics of the face are used for identification of humans. The detected face is compared with the dataset stored in the database using four-patch Local Binary Pattern Histogram with CNN.

1.1. Challenges of Face Recognition

- a. **Low Resolution:** Images with less than 16*16 resolutions are called low resolution images as the standard image must have minimum 16*16 resolution. Security cameras in super market, ATM cameras, CCTV cameras in streets and such small-scale standalone cameras can be found with these low-

resolution pictures. In these cameras only face region less than 16×16 gets captured as camera is not very close to the face and thus it captures only small part of human face. Much information is not provided by such low-resolution pictures and more information are lost. In recognizing the faces, it can be a big challenge.

- b. **Illumination:** One of the major problems of pattern matching in indoor as well as outdoor is illumination. According to the studies on Face Recognition, there are two problems to textual based illumination handling. Primarily, illumination normalization changes actual pixels of the face as result of increased contrast[3][4]. Besides, it increases false acceptance rate by imposing limit on the separation between classes.
- c. **Aging:** The characteristic values of face changes over a period of time and reflect as aging. After a long term of access, age invariants also become important for visual observation and retrieval of images. For different age groups, various face set over a period of time is checked for calculation of accuracy in Face Recognition system.
- d. **Poses:** Handling varying poses is one of the major challenges in Face Recognition technique[2][3][4]. The rotations cause the face image differences and that is larger than the interpersonal differences used in distinguishing identities. On the other hand, Face Recognition across pose has great potential in many applications where biometric technique can be implemented and utilized.
- e. **Occlusion:** In the Face Recognition system occlusion is the critical challenge[3]. If the full face is not visible in the input image due to moustache, beard, cap, mask or any accessories occlusion occurs. It is common in real world scenario. This blockage of face makes the automated Face Recognition process a challenging and difficult to solve.

1.2. Face Recognition Technology

Expressing face information completely by single face feature is difficult. In the application of Face Recognition, feature learning approaches based on deep learning have gained much attention. Here CNN is the most typical algorithm that could learn the features directly from the original image thereby avoids the complex feature extraction process. LBP is used in different Face Recognition applications and it is designed to extract different local features for the recognition of images. Within a gray-level image a string of bits is generated to encode pixel similarities within a small group of neighbouring pixels surrounding a given pixel. In the facial analysis this simple method may fast capture gray-level variation in pixels. In the conventional LBP method, gaining much improvement is difficult due to the factor that these methods are rotation variant and sensitive to noise. In the proposed system we tried to put the concept of four-patch LBP and CNN for improved accuracy in Face Recognition rate on different face databases like ORL, YALE, YOUTUE databases. Figure 1 shows the different stages of Face Recognition.

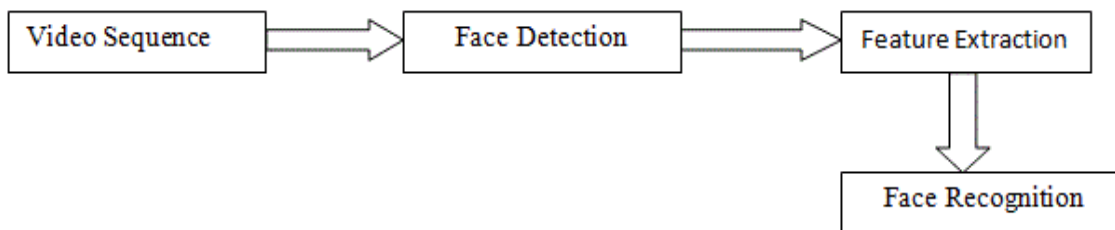


Figure 1. Stages of Face Recognition

The remaining sections of this paper are structured as follows. In section II related works are discussed in detail. The III section gives detailed explanation about techniques used in this work like Haar cascade classifiers, 4 patch LBP and CNN. The observations made by applying these techniques to different databases are presented in section IV.

2. Literature Survey

The main reason for the performance degradation of the Face Recognition technique is non frontal image of the face. When the taken image is not frontal image of the face, performance of most of the Face Recognition system reduces. To overcome this challenge in paper[5], a novel technique called local linear regression is applied to generate virtual frontal face from the image of non-frontal face. In this approach, detected non-frontal face is divided into smaller parts and then linear regression is applied to all of these to generate actual frontal face. Based on the training set, linear mapping function is estimated and used to create frontal face from the detected non frontal face. The effectiveness of this approach is validated on the PIE faces. The outcomes show that proposed approach is stronger when contrasted with eigen light field and 3D morphable model.

Most of the existing Face Recognition systems have several drawbacks like single sample problem and maximum reconstruction errors. These drawbacks reduce the Face Recognition rate. The Grey Wolf Optimizer based on Linear Collaboration discriminant Regression classifier is introduced in paper[6]. Here, recognition rate is improved by applying recognition technique of Grey Wolf Optimizer algorithm to Linear Collaboration discriminant Regression classifier. The efficient weighted values selected by the GWO algorithm are forwarded to LCDRC. The proposed approach reduces within class reconstruction error and increases between class reconstruction error of the training set. Authors have analyzed performance of the proposed work on ORL and YALE dataset. In this work, recognition accuracy of the introduced method is compared with Kernel Linear Regression Classification, Linear Regression Classifier, Linear Discriminant Classification and Linear Collaboration Discriminant Regression Classification. The outcomes show that the introduced method gives 3% and 6.5% enhancement with respect to ORL and YALE datasets respectively.

Nowadays, researchers are focusing on pose invariant Face Recognition as it is a difficult task. In paper[7], the impact of locally linear regression and fisher linear discriminant analysis (FLDA) on pose invariant Face Recognition is analyzed. Here, virtual front face is created using locally linear regression. Before giving created frontal face as input to FLDA, dimensionality is reduced by Principal Component Analysis to less than the difference between number of train examples and number of classes. This proposed work is analyzed using YALE dataset under various poses. The accuracy of 93.4% is obtained using proposed approach.

In paper[8], the mixture of PCA model along with fisherface method and SVD projections is used to get good accuracy. The proposed technique is analyzed on ORL dataset. The images of 40 people in different pose and illumination are considered for the analysis. For recognition purpose, the first dataset model is trained using fisherface to feature extraction and k-nearest neighbor to classification. Second model is trained using SVD based method. From the resulting sets of two trained models, a subset is obtained. All the labels are ranked based on the minimum distance when the subset contains greater than one label. And the label with least rank is selected. The experimental results show that recognition rate of PCA, fisherface and SVD is 80.7%, 98.4%, 96.6% respectively. By combining these three techniques the recognition rate is increased to 99.5%.

In paper[9], a hybrid method of Haar cascade and Eigenface methods is used for faster Face Recognition. This method can detect 55 faces in one detection process. With 91.67% accuracy, it is able to recognize multiple faces. Here, the process begins with preprocessing of the training data. Preprocessing includes grayscale conversion and normalization. Then Haar cascade algorithm is applied. The vital features are extracted from the preprocessed image using the combination of Eigenface and PCA. After the feature extraction, using Euclidian distance technique similarity distance is calculated between PCA_train and PCA_test data for Face Recognition.

The varying nature of the illumination, expression and pose reduce the accuracy of the Face Recognition techniques. Hence, lacking training samples are the main reason for the bottleneck of the Face Recognition. In paper[10], a new technique for generating different symmetrical face like left symmetrical face, right symmetrical face and mirror symmetrical face of the images is introduced. All the original and symmetrical training samples are used to generate score of test samples. The weighted score fusion technique is used to combine all these scores. Principal Component Analysis is then applied to classify these samples. The proposed method is applied to three datasets; YALE, ORL, FEI. This technique reduces the error rate in the Face Recognition by generating more samples for training the system.

In paper[11], a novel technique to increase the accuracy of Face Recognition in case of pose view face and expression input with database containing only frontal view of the face is introduced. With the help of Piecewise Affine Warping, proposed method converts the captured expression face to closed-mouth and opened-eyes face. Analysis is done on CMU Multi-PIE and CMU PIE dataset. Before applying the proposed approach, a facial image needs to be processed by active appearance model. In this work, output of the proposed technique that generates frontal view is given to PCA and LBP to check the accuracy of Face Recognition. It is observed that instead of giving detected image directly as input to PCA or LBP, giving output of the proposed work as input significantly enhances the Face Recognition accuracy.

A technique for improving Face Recognition when face occluded by sunglass and scarves is mentioned in paper[12]. The obstruction detection is done through improved support vector machine and principal component analysis. And recognition of remaining part is done through weighted local binary patterns. Initially face is divided into smaller disjoint patches. It is observed that by dividing the image of the face into six symmetric local patches better result can be obtained. The experimental outcome shows that the performance is more than

PCA, LR-PCA and LBP on non-occluded face and has recognition rate 94.45%. In case of face with scarf and glass also, proposed approach gives better result with recognition rate 91.05% and 74.83% respectively which is better than that of the PCA, LR-PCA and LBP.

The quality of the recorded footage may not be sufficient for recognizing the face when it is taken from far. Hence in paper[13], the performance of PCA is analyzed with nearest neighbor, bilinear and bicubic image enlargement techniques. The result is analyzed by sampling input image to six 512 x 786, 256 x 384, 128 x 192, 64 x 96, 32 x 48 and 16 x 24 resolutions by applying three image enlarging techniques separately. This approach is applied on SCFace database and the results show that Face Recognition accuracy of PCA is greater when images are enlarged by using nearest neighbor method as compared to other two image enlarging techniques.

Considering challenging problem of Face Recognition having variation in large pose and illumination in computer vision, in paper [14] a new learning-based face representation is proposed i.e., the face identity-preserving (FIP) features. Usage of this feature is variation in pose and illumination is removed significantly and variance in intra-identity is reduced by image of face reconstruction in canonical view. Negative effects from illumination and pose are eliminated using this canonical view reconstructed image of face as input by algorithm of conventional descriptors and learning. New strategy for training having two steps is proposed as there are large number of parameters. Least square dictionary-based initialization of parameter and by back propagating updating the parameter. Two experiment sets were conducted in which, descriptors of learning based and methods of state of the art gets compared in first set. And improved methods for recognition of classical face on image of face reconstruction in canonical view is demonstrated in second set. Improved performance is made by this property over traditional descriptors like LBP and Gabor. Including methods of 2D and 3D based, state of the art method outperformed by FIP features on Multi PIE database.

Classifying faces in training set to known identities is impossible and is a metric learning problem in Face Recognition. Large improvements to the Face Recognition (FR) have been done by deep convolution neural networks (CNNs). Dataset with more identities are trained on CNNs in FR and these datasets are too expensive to achieve state of the art performance. Paper [16], for studying discriminative feature of face, proposes a SqFace framework. Two loss functions jointly supervise the CNN model in this framework. They are auxiliary loss and chief classification loss like softmax loss, by combining together features of sequence and traditional identity dataset is learnt. Through sequence data full usage, deep face feature discrimination power is enhanced by employing label smoothing regularization and discriminative sequence agent. With single ResNet-64, experiments were conducted on datasets like labelled faces in the wild and you tube faces, where 99.83% LFW verification accuracy and YTF verification accuracy of 98.12% is achieved.

The convolutional neural network (CNN or ConvNet) is capable of extracting the features from the given input image and to recognize the face. By giving extracted features as input, one can further boost the performance of CNN. In this work, we have used four-patch LBP for feature extraction which will be then given to CNN for Face Recognition.

3. Methodology

3.1. Haar Cascade Classifier

The Haar Cascade classifiers is used in the present work for detecting the face. It identifies the presence or absence of face in the image. Haar Cascade is a face detection algorithm based on machine learning which is used to identify the object in a stationary image or video. Features are extracted from lot of positive (images with faces) and negative images (images without faces) to train the cascade function. For feature extraction Haar features shown in Figure 2 are used. The value obtained by subtracting total number of pixels in white rectangle from total number of pixels in black rectangle is considered as a feature. Depending on the training result the objects are detected from the new image. Each stage of the cascade classifier is a group of weak learners trained by boosting technique. Single weak learner is not capable to classify the images. Therefore, weighted average of the decisions made by each weak learner is considered as final classifier. Rather than applying all the features at a time on a window, small group of features are applied one by one. If the features in first group are not found in the current window then it is discarded. Otherwise, window is passed to the second group of features. When window passes all the group of features detector, the presence of face in the current window is reported.

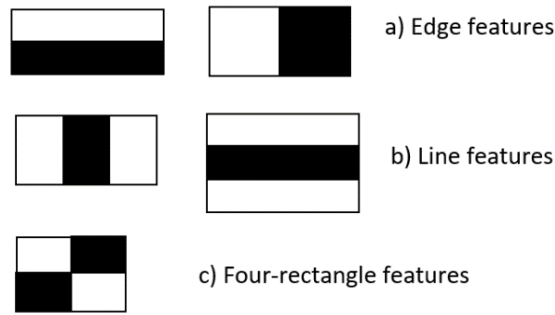


Figure 2. Types of Haar features

3.2. Four-patch LBP Codes

The Local Binary Pattern (LBP) is used to extract the features from the given face. To improve feature extraction, LBP is combined with Histograms of oriented gradients descriptor. LBPH divides the detected image into small sliding windows. These sliding windows are selected based on R (radius) and neighbours. In Figure 3, 3x3 sliding window is considered along with R and neighbours as 1 and 8 respectively. The new value for each neighbour is determined based on the threshold value. The central value of 3x3 matrix is considered as threshold value. The new binary value for the neighbour will be 1, if the original value is same or more than the threshold value and 0, if the original value is less than the threshold. These new binary values are concatenated in clockwise direction and the resultant binary value is converted into decimal value which is assigned to the centre of the sliding window. As a result, from the original image LBPH generates an image that highlights the facial features.

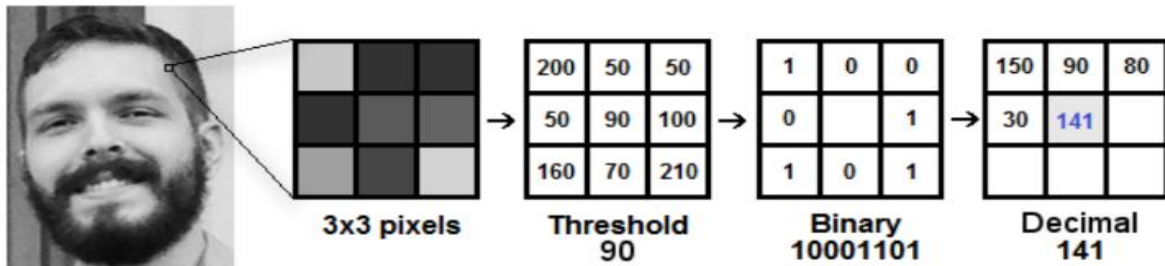


Figure 3. Local Binary Pattern

Now, the new image is divided into grids and a histogram representing intensity values is generated for each grid. Finally, as shown in Figure 4, all the histograms are concatenated to produce final histogram representing the features of the original face.

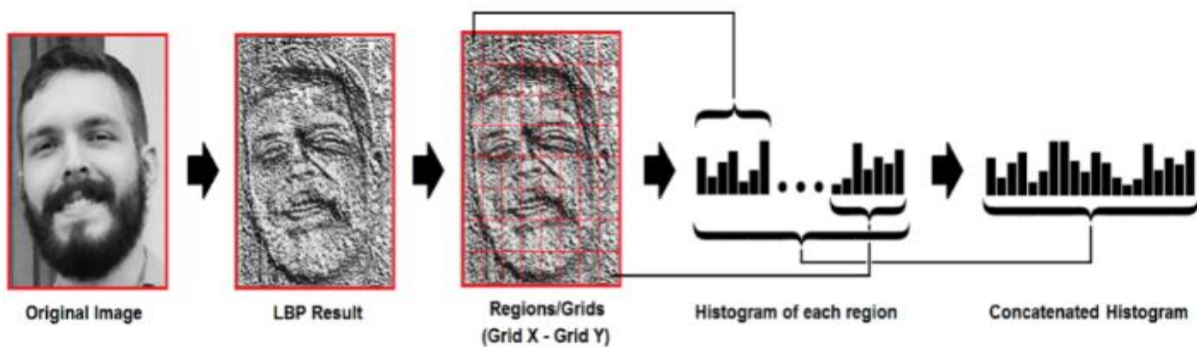


Figure 4. Histogram Concatenation

In this work, we have used four-patch LBP for feature extraction. The four-patch LBP code for a pixel is generated by considering two rings with considered pixel at the origin. The patches of size $n \times n$ spread out consistently on each ring as shown in Figure 5. Number of bits in the binary code will be half of the patches in each ring. The Four-patch LBP is generated by relating two centre symmetric patches in the outer ring with two centre symmetric patches in the inner ring located R patches away along the circle. As name indicates, based on the four patches it generates binary code for the pixel and leads to the more accurate result as compared to LBP and three-patch LBP. For example, in the Figure 5, four patch LBP can be given by

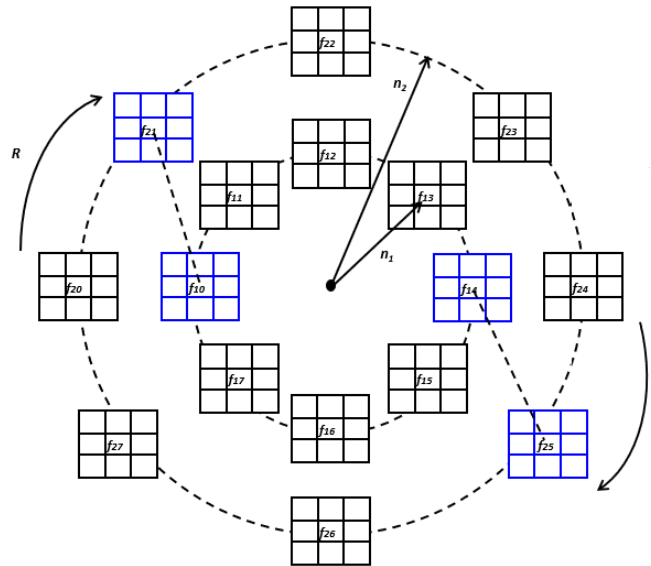


Figure 5. The four patch LBP code

$$4LBP(P_c) = D(f_{21}-f_{10}) - D(f_{25}-f_{14}) * 2^0 D(f_{22}-f_{11}) - D(f_{26}-f_{15}) * 2^1 D(f_{23}-f_{12}) - D(f_{27}-f_{14}) * 2^2 D(f_{24}-f_{13}) - D(f_{28}-f_{14}) * 2^3$$

The formal definition of four-patch LBP is given by,

$$4PLBP_{n_1, n_2, P, x, R}(\rho) = \sum_j^{P/2} S \left(\begin{matrix} w(f_{1,j}, f_{2, j+R \bmod P}) \\ -w(f_{1, j+P/2}, f_{2, j+P/2+R \bmod P}) \end{matrix} \right) 2^j$$

3.3. Face Recognition using Convolutional Neural Network

Face Recognition is a problem of multi class classification and in many methods by softmax loss, CNN models are supervised. It is proved by many methods that on some benchmark dataset, CNN outperform in FR. In deep learning, one of the represented network structures is CNN. In the areas like recognition of image and speech analysis CNN has become a hotspot [15]. In an image data space, CNN models are ubiquitous which works well for classification of images, recognition of image, detection of object etc. Reduced network model complexity and number of weights are achieved by CNN as it has weight sharing network structure. Further, traditional recognition algorithm’s feature extraction complexity and reconstruction of data are avoided as image is directly given as input to network. To recognize shapes of two-dimensional image which are invariant to tilt or other deformation forms, a multilayer product network is designed.

Local area perception, spatial sampling and weight sharing are three CNN characteristics which are important frame of idea. Sampling layer connected by convoluted layer and convoluted layer in connection to sampling layer are used in alternate set by CNN. Features extracted by convolution layer are combined later to form abstract feature and finally object characteristics of image description is formed. Fully connected layer can also be followed for CNN and these are the convolutional neural networks basic idea, whereas multiple versions also exists. DeepLearnToolbox is referred here. Structure diagram of CNN is shown in Figure 6.

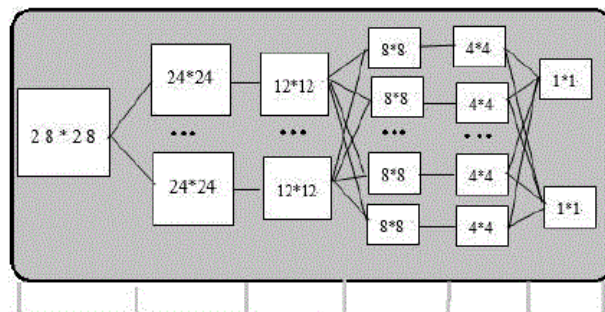


Figure 6. Input: 28*28, two convolutional layers, two pooled layers and output layers

The structure of neural network in shown figure has 28×28 input layer, 2 convolutions, 2 pooling layers and output layer which is fully connected. Six 24×24 characteristic matrix is generated when initial input image of 28×28 is convoluted to first layer and through this first layer six 12×12 matrix is generated. With the second layer of convolution, 8×8 feature matrices is generated and then sampling of second layer convolution features generates 12×4 feature matrices. And output vector of 2×1 is generated when 2 neurons are fully connected by last twelve 4×4 feature matrix. Recognition of face is realized by building good CNN in MATLAB. Flow chart of CNN based face recognition is shown in Figure 7.

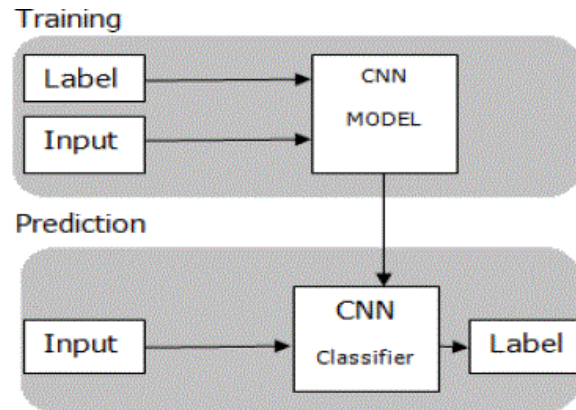


Figure 7. CNN model for identifying face

First classifier is generated by training the CNN using training data set. Later with test dataset result of classification is tested.

4. Experimental Results

In this part, results of our experiment are presented in detail. The implementation is done with 3.2GHz Pentium Core i5 processor using MATLAB version 7.2. The effectiveness of the proposed algorithm is evaluated on the reputed datasets like UPC Face Database Face, ORL Face Database, Yale B, YouTube Face Video Database (YTF) and Labelled Faces in Wild Database (LFW). These datasets are described below in detail.

4.1. UPC Video Face Database (GTAV Face Database)

Against strong variation in pose and illumination, for testing robustness of Face Recognition algorithms this database is created. There are total 44 persons included in the database. Corresponding to different pose view (0° , $\pm 30^\circ$, $\pm 45^\circ$, $\pm 60^\circ$ and $\pm 90^\circ$) with three different illuminations (strong light source from an angle of 45° , natural light or environment and finally an almost frontal mid-strong light source) each person's 27 pictures are included. With different occlusions and facial expression variations, further in addition at least 10 more frontal view pictures are included. Images are in BMP format with resolution 240×320 . The Figure 8 show sample from UPC dataset.



Figure 8. Sample video frames from the UPC database

4.2. You Tube Face (YTF) Database

In this data set, the smallest clip has 48 frames whereas, lengthiest clip has 6070 frames. On an average 2.15 videos of each person are available in YTF database. Figure 9 shows a sample set from YTF. In total it contains 3425 videos of 1595 persons.



Figure 9. Sample video frames from the YTF database

4.3. Labelled Faces in Wild (LFW)

The problem of unconstrained Face Recognition can be studied through the LFW database of face photographers. There are 13000 images of faces (two or more different photos of 1680 people) taken from the web with name of the person as label. Figure 10 shows the sample LFW Dataset.



Figure 10. Sample images from the LFW dataset

4.4. Yale B Database

This database has 5760 facial images of 10 people. Here, each person's 576 different images with different facial expressions are considered. Each face is cropped into a size of 320×320 ". Figure 11 shows a sample Face set from the Yale B database.



Figure 11. Sample images from the Yale B Face database

4.5. ORL Database

This database includes 400 face images of 40 persons with 10 different expressions. The face images were taken under different light conditions and with different facial expressions. The Figure 12 shows sample images.



Figure 12. Sample images of ORL database

Table 1 shows the Databases used for our experiments. The table indicates the name of the Database, Identifies the Number of Images per Identities, Image Size and Image type.

Table 1. The Databases and the data samples

Datasets	Identities	Images	Images identities	Image size	Image type
ORL	40	400	10	92 × 112	JPEG
UPC(GTAV Face)	44	704	16	240 × 320	BMP
YouTube face (YTF)	50	700	14	320 × 240	JPEG
Labelled faces in the wild (LFW)	50	700	14	250 × 250	JPEG
Yale B	10	5760	15	320 × 320	JPEG

4.6. Face Recognition Performance Evaluation with Confusion matrix

In this work, different statistical measures are used to evaluate the performance of the system. The result may be positive (authorized person) or negative (unauthorized person). The following settings are made for the considered statistical measures.

- TP – Perfect recognition of an authoritative individual
- FP – Authoritative recognition of an Unreliable individual.
- FN – Unreliable recognition of an Authoritative individual.
- TN – Perfect recognition of an Unreliable individual.

Using the above values, the statistical measures are calculated. Different statistical measures are calculated as follows,

Sensitivity (Recall)- is calculated by:

$$Sensitivity = \frac{TP}{TP + FN}$$

Sensitivity (namely True_Positive_Rate, Re-call else Probability_of_Detection) process the amount of constructiveness so as to accurately recognize.

a. Specificity- is calculated by:

$$Specificity = \frac{TN}{TN + FP}$$

Specificity (namely True_Negative_Rate) determines the ration of downbeats so as to accurately recognize.

b. Positive predictive value (PPV) - is calculated by:

$$PPV = \frac{TP}{TP + FP}$$

Positive_Predictive_Value is the ratio of constructiveness to facilitate the occurrence of a form.

c. Negative predictive value (NPV) - is calculated by:

$$NPV = \frac{TN}{TN + FN}$$

Negative_Predictive_Value is the ratio the ration of downbeats which matches up to the nonexistence of the form.

d. Accuracy- is calculated by:

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

The accuracy is a definite percent of perfectly categorized occurrence.

e. False Positive Rate(FPR)- is calculated by:

$$FPR = \frac{FP}{FP + TN}$$

f. False Discovery Rate(FDR)- is calculated by:

$$FDR = 1 - PPV = \frac{FP}{FP + TP}$$

Mathew Correlation Coefficient(MCC)- is calculated by:

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

MCC computes the Mathews Correlation Coefficient between observed and predicted binary classifications.

For the experiment, four different datasets are formed with different size: Dataset1, Dataset2, Dataset3, Dataset4. Dataset1 contains the 10 frames of the known individual, Dataset 2 contains 8 frames of the same person, Dataset 3 contains 6 frames of the same person and Dataset 4 contains 4 frames of the same person. All

the individuals are the known personalities from YTF. Table 2 shows the Result of the experiment. From the experiment Maximum Accuracy of 94.00% and an Average Accuracy of 92.3% is reported.

Table 2. Recognition result with YTF Video Face

Statistical Measures	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Average
Sensitivity	96.00	94.00	94.00	88.00	93.0
Specificity	92.00	94.00	92.00	88.00	91.5
Precision	92.31	94.00	92.16	88.00	91.6
Negative Predictive Value	95.83	94.00	93.88	88.00	92.9
False Positive Rate	08.00	06.00	08.00	08.00	7.5
False Discovery Rate	07.69	06.00	07.84	08.00	7.3
False Negative Rate	04.00	06.00	06.00	08.00	6.00
Accuracy	94.00	94.00	93.00	88.00	92.3
MCC	88.07	88.00	86.07	76.00	85

Next, the Experiment is conducted with ORL and Yale B Face Sets. Both the Databases contain face images of the known individuals. For experiment purpose seven Datasets are formed by selecting the image samples from both ORL and Yale B Face sets separately. The experiment result is shown in the Table 3. From the experiment it is clear that the ORL Database shows a maximum performance accuracy of 94.1% with eight samples and Yale B shows a maximum accuracy of 86.82% with eight samples.

Table 3. FR with ORL database and YALE B Database

Database	Accuracy (%)						
	Two training samples	Three training samples	Four training samples	Five training samples	Six training samples	Seven training samples	Eight training samples
ORL	85.1	87.32	89.61	90.89	92.14	93.89	94.01
Yale B	67.86	72.01	74.50	78.59	81.80	84.25	86.82

Table 4 and the Figure 13 shows the statistical measures from the experiment conducted on GTAV (UPC) Video Database. This database contains different Poses of the individuals. The result shows a maximum accuracy of 95% with four-patch LBP and 93.5% with 1 patch LBP-CNN. The Figure clearly shows the performance of both the methods.

Table 4. Recognition Rates (%) on Different Poses on UPC Video database

Pose angle/Methods	0°	+30°	-30°	+45°	-45°	+60°	-60°	Avg
4-Patch LBP-CNN	95.00	89.50	89.50	85.00	85.00	83.16	83.16	82.11
LBP-CNN	93.50	83.16	83.33	83.16	80.00	77.00	78.50	82.43

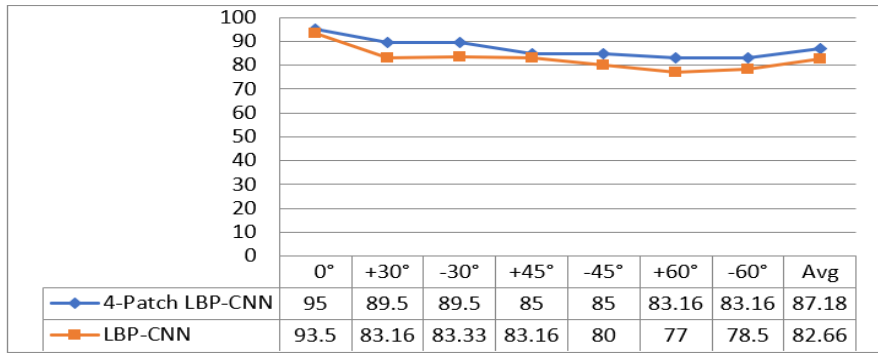


Figure 13. Performance Analysis

Finally, the performance of the proposed four-patch LBP-CNN model is compared with other recent techniques. All the selected models are based on Deep Learning techniques. The following table, Table 6 shows the result of proposed and state-of-Art techniques. Here the proposed four-patch with CNN shows a maximum accuracy of 94.0%.

Table 5. Computation speed in Seconds

Experiment No.	LBP	4-Patch LBP
1	04.5000	04.67961
2	06.6456	05.7800
3	06.4567	06.5018
4	06.7654	06.8294
5	06.9780	07.1849
6	06.7653	06.7121
7	05.4887	05.79376
8	05.9856	07.51872
9	04.9399	05.74058
10	06.8909	07.79030

Table 6. Verification on YouTube Faces DB

No	Method	YTF
1	Proposed method	94.0%
2	SeqFace, 1 ResNet-64	98.12%
3	ArcFace + MSIMV2 + R100	98.02%
4	CosFace	97.60%
5	Light CNN-29	95.54%
6	Git Loss	95.30%
7	FaceNet	95.12%
8	SphereFace	95.00%
9	DeepId2+	93.20%
10	3DMM face shape parameters + CNN	88.80%

The experiment is conducted to record the recognition speed. Ten separate experiments are conducted and the speed of the recognition is recorded in the Table 5. The speed of four-Patch LBP with CNN is also compared with the LBP-CNN model.

5. Conclusion

The proposed Face Recognition method is implemented with the help of four-patch LBP and CNN. From the result analysis it is clear that the LBP is an illumination invariant method and hence, it is very useful in recognizing the faces with varying illumination conditions. The training of the CNN is carried out using the standard datasets such as YTF, ORL, Yale B, and GTAV. The outcome of the experiments conducted demonstrates that the proposed technique is proficient in recognizing the Faces with least computation speed. By comparing the accuracy, sensitivity, specificity and other statistical measures, it is clear that the proposed method results good performance. It is clear that the suggested recognition model obtains a maximum accuracy of 94.01% with ORL datasets, 86.82 with Yale B, 94% with You Tube Database. Pose angle variation is one of the major problems in FR, in order to overcome this problem, we have selected the datasets from GTAV(UPC) Database. The experiment conducted with UPC video frames shows a maximum accuracy of 95% with 4-patch LBP and 93.2% with 1-patch LBP features respectively. A minimum computation speed of 4.67 Seconds including training time, is observed from the experiment with UPC dataset which adds more strength to the proposed technique.

References

1. Shreekumar, T., & Karunakara, K. (2019). Face Pose and illumination Normalization for Unconstraint Face Recognition from Direct Interview Videos. *International Journal of Recent Technology and Engineering (TM)*, 7(6S4), 59-68.
2. Shreekumar, T., & Karunakara, K. (2019). Face Pose and Blur Normalization for Unconstraint Face Recognition from Video/Still Images. *International Journal of Innovative Computing and Applications, Inderscience*, ISSN: 1751-648X
3. T. Shreekumar, K. Karunakara. (2019). Identifying the Faces from Poor Quality Image / Video. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 8(12), 1346-1353.
4. Shreekumar, T., & Karunakara, K. (2018). A Video Face Recognition System with Aid of Support Vector Machine and Particle Swarm Optimization (PSO-SVM). *Journal of Advanced Research in Dynamical and Control Systems (JARDCS)*, 05, 496-507
5. Chai, X., Shan, S., Chen, X., & Gao, W. (2007). Locally linear regression for pose-invariant face recognition. *IEEE Transactions on image processing*, 16(7), 1716-1725. <https://doi.org/10.1109/TIP.2007.899195>
6. Hosgurmath, S., & Mallappa, V.V. (2019). Grey Wolf Optimizer with Linear Collaborative Discriminant Regression Classification based Face Recognition. *International Journal of Intelligent Engineering and Systems*, 12(2), 202-210.
7. J. Shermina. Impact of locally linear regression and Fisher Linear discriminant analysis in pose invariant face recognition. *Int. J. Comput. Sci. Netw. Secur.*, 10 (10) (2010), pp. 106-110
8. Shermina, J. (2010). Impact of locally linear regression and fisher linear discriminant analysis in pose invariant face recognition. *International Journal of Computer Science and Network Security*, 10(10), 111-115.
9. Dhamija, J., Choudhury, T., Kumar, P., & Rathore, Y.S. (2017). An Advancement towards Efficient Face Recognition Using Live Video Feed:" For the Future". In 2017 3rd *International Conference on Computational Intelligence and Networks (CINE)*, 53-56. <https://doi.org/10.1109/CINE.2017.21>
10. Mantoro, T., & Ayu, M.A. (2018). Multi-faces recognition process using Haar cascades and eigenface methods. In 2018 6th International Conference on Multimedia Computing and Systems (ICMCS), 1-5. <https://doi.org/10.1109/ICMCS.2018.8525935>
11. Siddiqui, M. Y. F. (2015, September). Face recognition using original and symmetrical face images. In 2015 1st International Conference on Next Generation Computing Technologies (NGCT), 898-902. <https://doi.org/10.1109/NGCT.2015.7375249>
12. Petpairote, C., Madarasmi, S., & Chamnongthai, K. (2017). A pose and expression face recognition method using transformation based on single face neutral reference. 2017 *Global Wireless Summit (GWS)*, Cape Town, 123-126. <https://doi.org/10.1109/GWS.2017.8300485>
13. Chen, Z., Xu, T., & Han, Z. (2011). Occluded face recognition based on the improved SVM and block weighted LBP. In 2011 International Conference on Image Analysis and Signal Processing, 118-122. <http://doi.org/10.1109/IASP.2011.6109010>

14. Halidu, M.K., Bagheri-Zadeh, P., Sheikh-Akbari, A., & Behringer, R. (2019). Pca in the context of face recognition with the image enlargement techniques. In 2019 8th *Mediterranean Conference on Embedded Computing (MECO)*, 1-5. <https://doi.org/10.1109/MECO.2019.8760162>
15. Zhu, Z., Luo, P., Wang, X., & Tang, X. (2013). Deep learning identity-preserving face space. In *Proceedings of the IEEE International Conference on Computer Vision*, 113-120. <https://doi.org/10.1109/ICCV.2013.21>
16. Hu, W., Huang, Y., Zhang, F., Li, R., Li, W., & Yuan, G. (2018). SeqFace: make full use of sequence information for face recognition. *arXiv preprint arXiv:1803.06524*.
17. Zhang, H., Qu, Z., Yuan, L., & Li, G. (2017). A face recognition method based on LBP feature for CNN. In 2017 IEEE 2nd *Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 544-547.