



Automatic Face Recognition Using Enhanced Firefly Optimization Algorithm and Deep Belief Network

Prakash Annamalai^{1*}

¹*Tagore Engineering College, Department of Information Technology, India*

* Corresponding author's Email: prakash1712@yahoo.com

Abstract: Face recognition is an emerging research area in biometric identification systems, which attracted many researchers in pattern recognition and computer vision. Since, the human facial images are high dimensional in nature, so a new optimization algorithm was proposed in this paper to reduce the dimension of facial images. At first, the input facial images were collected from ORL, YALE, and Face Semantic SEGmentation (FASSEG) databases. Then, feature extraction was then performed by using Local Ternary Pattern (LTP) and Binary Robust Invariant Scalable Key points (BRISK) to extract the features from the face images. In addition, an enhanced fire-fly optimizer was used to reject the irrelevant features or to reduce the dimension of the extracted features. In enhanced fire-fly optimization algorithm, Chi square distance measure was used to find the distance between the fire-flies and also it uses only a limited number of feature values for representing the data that effectively reduces the “curse of dimensionality” concern. Finally, Deep Belief Network (DBN) was used to classify the individual’s facial images. The experimental outcome showed that the proposed work improved recognition accuracy up to 7-20% compared to the existing work.

Keywords: Binary robust invariant scalable key points, Deep belief network, Enhanced fire-fly optimizer, Face recognition, Local ternary pattern.

1. Introduction

In recent years, face recognition is extensively studied due to its applications in several fields like human computer interaction, internet network communication, computer entertainment, artificial intelligence, e-commerce security, and security monitoring [1-3]. Although, several researchers have achieved significant advances in face recognition system, still this application have many challenges like intra-class variations in illumination, expression, noise, pose and occlusion [4, 5]. So, robust and discriminant feature generation have become a problem in face recognition [6]. In addition, the dimensionality of collected human face images is often very high, which leads to high computational complexity and a curse of dimensionality issue [7, 8]. Currently, various dimensionality reduction techniques are developed in semi-supervised, unsupervised, and supervised

circumstances like Volterra kernels direct discriminant analysis [9], principal component analysis [10], linear discriminant analysis [10], etc. Additionally, the prior approaches classify the individuals on the basis of hand designed manner that consumes more time for individual classification [11, 12].

To address this issue, a new system is proposed for achieving good performance in face recognition. Initially, the input facial images were collected from the ORL, YALE, and FASSEG databases and then histogram equalization was applied for denoising the collected images. The histogram equalization calculates the histograms from the images to redistribute the lightness pixel values, which significantly improves the contrast and edges of the images. Then, feature extraction was performed by utilizing BRISK, and LTP for extracting the feature vectors from the denoised images. After attaining the hybrid feature values, enhanced fire-fly

optimizer was applied to reduce the dimension of the extracted features. In enhanced fire-fly optimizer, the Chi square distance was utilized instead of Cartesian distance to find the distance between the fire-flies. Hence, the Chi square distance utilizes only a limited number of features to represent the data that effectively reduces the “curse of dimensionality” concern. Then, the output of feature selection was given as the input for DBN classifier for classifying the facial images. At last, the proposed work performance was compared with an existing work in light of accuracy, precision, and recall.

This research paper is prepared as follows. In section 2, numerous research papers on face recognition topic are reviewed. The detailed explanation about the proposed work is given in the section 3. In addition, section 4 illustrates about the quantitative and comparative analysis of the proposed work. The conclusion is made in section 5.

2. Literature survey

Several approaches are developed by the researchers in face recognition topic. In this section, a brief review of some important contributions to the existing literature is presented.

L. Zhou, W. Li, Y. Du, B. Lei, and S. Liang [13] developed an illumination invariant local binary descriptor learning methodology for face recognition. Usually, the developed methodology utilizes the rigid sign function for binarization instead of data distribution. Initially, the developed method calculates the dynamic thresholds, which includes the information of illumination variation to extract the nonlinear multi-layer contrast features. In this research, exponential discriminant analysis was used as a pre-processing method that significantly enhance the discriminative ability of the face image by enlarging the margin between dissimilar classes relative to the same class. In addition, the adaptive fuzzy fusion system was used for incorporating the recognition results for multi-scale features space. From the experimental simulation, the developed methodology attained better consequence related to the existing works in light of recognition accuracy. The developed methodology performance was highly degraded in the conditions like illumination conditions and facial variations.

A. Ouyang, Y. Liu, S. Pei, X. Peng, M. He, and Q. Wang [14] presented a hybrid approach for face recognition by combining Probabilistic Neural Networks (PNNs) and Improved Kernel Linear Discriminant Analysis (IKLDA). Initially, the dimension of the extracted features was reduced in

order to retain the relevant information. Then, PNN methodology was adopted for solving the face recognition problems. The developed method (IKLDA+PNN) not only improves the computing efficiency, and also its precision value. In this literature, the experimental investigation was conducted on the AR, ORL, and YALE databases, which comprises of a wide variety of facial details, and expressions. The experimental result showed that the developed methodology (IKLDA+PNN) attained better performance by means of standard deviation and accuracy. If the number of collected images were low, the developed hybrid approach significantly lessens the recognition performance.

Q. Liu, C. Wang, and X.Y. Jing [15] developed a new nonlinear feature extraction approach (dual multi-kernel discriminant analysis) for colour face recognition. At first, a kernel selection method was designed for selecting the optimum kernel mapping function for each colour component of the facial images. Then, a colour space selection model was developed for choosing the appropriate space, and then map the dissimilar colour components of the facial images into dissimilar high dimensional kernel spaces. At last, discriminant analysis and multi kernel learning were applied not only within each component but also between dissimilar components. The performance of the developed approach was tested on face recognition grand challenge version 2 and Labelled Faces in the Wilds (LFW) databases. From the experimental outcome, the developed approach outperforms the existing approaches in terms of recognition rate. In this work, the training and testing process requires manual intervention, which needed to be automated.

C.H. Yoo, S.W. Kim, J.Y. Jung, and S.J. Ko [16] developed a new feature extraction approach by utilizing the high dimensional feature space for enhancing the discriminative ability for face recognition. At first, the local binary pattern was decomposed into several bit planes that consists of scale specific directional information of the facial images. Each bit plane includes inherent local structure of the facial images and also the illumination robust characteristics. In addition, a supervised dimension reduction approach (orthogonal linear discriminant analysis) was used to lessen the computational complexity, when preserving the incorporated local structural information. The extensive experiments were conducted and validated the efficiency of the developed approach on AR, LFW and Extended Yale B datasets. In contrast, the developed methodology failed to achieve better recognition accuracy in a large facial dataset.

I.M. Revina, and W.S. Emmanuel, [17] developed a new system for facial expression recognition that combines whale grasshopper optimization approach and multi support vector neural network. Initially, scale invariant feature transform and scatter local directional pattern were used to extract the features from the collected facial images. The extracted feature values were classified by applying the developed classifier for recognizing the facial expressions. In the experimental phase, the developed system was tested on Japanese female facial expression database and Cohn-Kanade AU-Coded Expression dataset. The developed system outperforms the existing works in light of recognition accuracy. By utilizing low level features in high level network, the semantic gap was maximized and it leads to poor recognition rate.

In order to address the above-mentioned concerns, a new facial recognition system is proposed in this paper.

3. Proposed system

In recent years, facial recognition is an emerging research area in the fields of artificial intelligence and pattern recognition. Hence, the face recognition is a challenging topic, because the real world facial images are formed with the interaction of multiple factors such as facial rotation, illumination conditions, background interferences, etc. In order to develop an effective facial recognition system, some of the problems need to be resolved, particularly on feature extraction and classification.

While classifying, the individual facial images are converted into vectors that are very high dimensional in nature. In order to lessen the “curse of dimensionality” concern, a new facial recognition system is proposed in this research paper. The proposed facial recognition system comprises of five steps such as image collection, image denoising, feature extraction, optimization, and classification. The working process of proposed facial recognition system is indicated in Fig. 1. The detailed explanation about the proposed facial recognition system is given below.

3.1 Image pre-processing

At first, the input face images are collected from ORL, YALE, and FASSEG datasets. After collecting the facial images, histogram equalization is used for pre-processing the images. It is a technique to adjust facial image intensities for enhancing the image contrast.

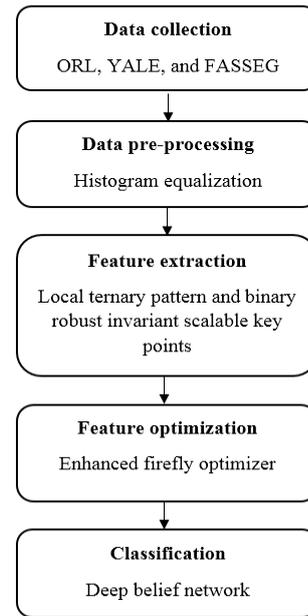


Figure. 1 Work flow of proposed system

Let I be a facial image that is represented by matrix integer pixel intensities m ranging from 0 to $L - 1$. The histogram equalized facial image H is defined in Eq. (1).

$$H_{i,j} = \text{floor} (L - 1) \sum_{n=0}^{I(i,j)} p_n \quad (1)$$

Where, L is indicated as possible intensity value of 256, p is denoted as normalized histogram of I with a bin for every possible intensity, and $\text{floor}()$ rounds down the nearest integers that are equivalent to transform the pixel intensities k , and it is expressed in Eq. (2).

$$T(k) = \text{floor}(L - 1) \sum_{n=0}^k p_n \quad (2)$$

In transformation, the intensities I and H are considered as continuous random variables of $X = Y$ on $[0, L - 1]$ with Y , which is mathematically defined in Eq. (3).

$$Y = T(X) = (L - 1) \int_0^x p_X(x) dx \quad (3)$$

Where, p_X is represented as probability density function of I , T is represented as cumulative distributive function of X multiplied by $(L - 1)$, and T is differentiable and invertible. It shows that Y is determined by $T(X)$, which is uniformly distributed on $[0, L - 1]$ namely $p_Y(y) = \frac{1}{L-1}$ as expressed in the Eqs. (4), (5), and (6) [18]. The original and enhanced facial images are graphically denoted in Fig. 2.

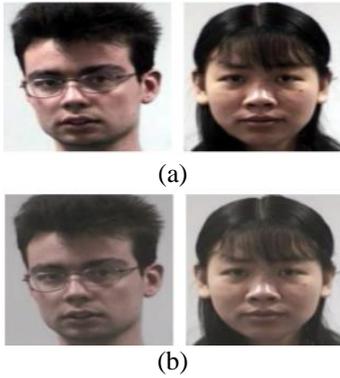


Figure. 2 (a) original image and (b) denoised image

$$\int_0^y p_Y(z) dz = \frac{1}{L-1} = \int_0^{T^{-1}(y)} p_X(w) dw \quad (4)$$

$$\begin{aligned} \frac{d}{dy} \left(\int_0^y p_Y(z) dz \right) &= \\ p_Y(y) &= p_X(T^{-1}(y)) \frac{d}{dy} (T^{-1}(y)) \end{aligned} \quad (5)$$

$$\begin{aligned} \frac{dT}{dx} \Big|_{x=T^{-1}(y)} \frac{d}{dy} (T^{-1}(y)) &= (L - \\ 1) p_X(T^{-1}(y)) \frac{d}{dy} (T^{-1}(y)) &= 1 \end{aligned} \quad (6)$$

3.2 Feature extraction

After denoising the facial images, feature extraction is accomplished for extracting the feature values. Here, hybrid feature descriptors are applied to extract the feature information from the denoised facial images. The hybrid feature descriptors include BRISK, and LTP. A brief description about hybrid feature extraction is detailed below.

3.2.1. Binary robust invariant scalable key points

BRISK is a texture descriptor that attains a significant quality of matching with limited computation time and also generate a valuable key-points from a face image. Generally, BRISK descriptor applies a symmetric sampling pattern over sample point of smooth pixels in feature descriptor. The image intensity is represented as i_x and then apply Gaussian smoothing with standard deviation σ_x that is equivalent to the distance between the points and circles. The key-point k in a facial image is patterned according to its scaling and position, and the sampling-point pairs are denoted as (i_x, i_y) . Correspondingly, the intensity of smoothed values of points is represented as $S(i_x, \sigma_x)$ and $S(i_y, \sigma_y)$ that helps to find the local gradients [19]. The local gradient $G(i_x, i_y)$ is mathematically denoted in Eq. (7).

$$G(i_x, i_y) = (i_y - i_x) \times \frac{S(i_y, \sigma_y) - S(i_x, \sigma_x)}{\|i_y - i_x\|^2} \quad (7)$$

Then, assume the set A of sampling point pairs as defined in Eq. (8).

$$\begin{aligned} A &= \{(i_x, i_y) \in \mathbb{R}^2 \times \mathbb{R}^2 | x < \\ N \wedge y < x \wedge x, y \in N\} \end{aligned} \quad (8)$$

Where, N is indicated as number of sampling point pairs. Then, partition the pixel pairs into two sub-sets; short distance pairs d_1 and long distance pairs d_2 that is mathematically indicated in the Eqs. (9) and (10).

$$d_1 = \{(i_x, i_y) \in A | \|i_y - i_x\| < \delta_{max}\} \subseteq A \quad (9)$$

$$d_2 = \{(i_x, i_y) \in A | \|i_y - i_x\| < \delta_{min}\} \subseteq A \quad (10)$$

The threshold distance is set as $\delta_{max} = 9.75t$ and $\delta_{min} = 13.67t$ (where t is the scale of k). Hence, the point pairs are iterated using L for finding the complete pattern direction of the key points k that is mathematically represented in Eq. (11).

$$G = \begin{pmatrix} G_x \\ G_y \end{pmatrix} = \frac{1}{L} \times \sum_{(i_x, i_y) \in L} G(i_x, i_y) \quad (11)$$

The sampling pattern rotation of orientation is represented as $\alpha = \arctan2(G_y, G_x)$ of the key-point. The binary descriptor b_k is generated by using short distance paring and each bit b_k is calculated from a pair in F . In this scenario, the descriptor length is 512 bits, which is fixed by performing the short distance intensity at each binary feature vectors v as stated in Eq. (12).

$$v = \begin{cases} 1 & S(i_y^\alpha, \sigma_y) > S(i_x^\alpha, \sigma_x) \\ 0 & \text{Other wise} \end{cases} \forall (i_x^\alpha, i_y^\alpha) \in F \quad (12)$$

3.2.2. Local ternary pattern

LTP is an extension of Local Binary Pattern (LBP) that utilizes a thresholding constant for thresholding the pixels into three values. Let us consider, k as a thresholding constant, p as a neighboring pixels, and c as a center pixel value. The result of the thresholding is determined by using the Eq. (13).

$$LTP(p, c, k) = \begin{cases} 1 & \text{if } p > c + k \\ 0 & \text{if } p > c - k \text{ and } p < c + k \\ -1 & \text{if } p < c - k \end{cases} \quad (13)$$

In this scenario, every threshold pixel has one of the three values and the neighbouring pixels are combined after thresholding into a ternary pattern. By calculating the histogram of ternary values results in a large range, so the ternary pattern is split into two binary patterns. The histograms are concatenated for generating a descriptor (double the size of LBP), which is very successful in the application of facial recognition. The basic idea of LTP is to transform the intensity space to order space, where the order of neighbouring pixels is utilized for creating a monotonic change illumination invariant code for every image [20].

3.3 Feature optimization

After feature extraction, optimization is accomplished for reducing the dimension of the extracted features. In recent periods, numerous optimization algorithms are developed for feature optimization. In that, fire-fly optimizer is superior in dealing with global optimization concerns. The fire-fly optimization algorithm is established by Xin-She Yang at Cambridge University based on the behaviour of fire-fly and flashing patterns. Generally, fire-fly optimizer contains three idealized rules such as attractiveness, unisex and brightness. Usually, all fire-fly insects are unisex, so it is attracted to other fire-flies on the basis of their sex. The brightness of the fire-flies is evaluated by the land-scape of the objective function value. In addition, the attractiveness is directly proportional to the brightness of a fire-fly, whereas the less bright fire-flies are attracted by the brighter fire-flies.

The attractiveness between the two fire-flies is automatically decreased, if the distance between the fire-flies is increased. The two major problems in fire-fly optimization algorithm are formulating the amount of absorbency and difference in the light intensity. For simplicity purpose, consider fire-fly absorbency β with luminosity that depends on the target functions. Since, the fire-fly absorbency is proportional to the light intensity of adjacent fire-fly. The absorbency from fire-fly β is mathematically defined in Eq. (14).

$$\beta(r) = \beta_0 e^{-\gamma r^2} \quad (14)$$

Where, β_0 is represented as attractiveness at $r = 0$ and γ is represented as coefficient of light

absorption at the source. The distance between the fire-flies j and i in the points x_j and x_i is mathematically represented in the Eqs. (15) and (16) on the basis of Chi square distance measure.

$$r_{ij} = \|x_i - x_j\| \quad (15)$$

$$r_{ij} = \sqrt{\frac{1}{2} \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{[x(i,j) - y(i,j)]^2}{x(i,j) + y(i,j)}} \quad (16)$$

Where, L is indicated as total number of gray levels, x_i and x_j are denoted as spatial coordinates of the fire-flies i and j , and $x_{i,k}$ is represented as the k portion of spatial coordination x_i of fire-fly i in the two dimensional state. Usually, Cartesian distance is used for finding the distance between the fire-flies j and i . In enhanced algorithm, Chi square distance is used to identify the distance between the fire-flies j and i , because it consumes only limited time for deciding the distances between the fire-flies j and i that is enough for achieving accurate neighbourhood selection for better recognition. The fire-fly movement and absorption of j moves luminously, which is determined by using the Eq. (17).

$$x_i = x_i + \beta_0 e^{-\gamma r_{ij}^2} (x_i - x_j) + \alpha (rand - \frac{1}{2}) \quad (17)$$

Where, the second term $\beta_0 e^{-\gamma r_{ij}^2} (x_i - x_j)$ is denoted as attractiveness, third term $\alpha (rand - \frac{1}{2})$ is represented as randomization with a parameter α , and $rand$ is stated as random number generator, which ranges from $[0, 1]$ interval.

In most of the cases, β_0 is equal to one and $\alpha \in [0, 1]$. The parameter γ determines the variations in attractiveness and also its objective function value is essential in evaluating the speed of convergence and fire-fly behaviour. Theoretically, the parameter γ is evaluated on the basis of absorbency changes that ranges within $[0, \infty]$. The pseudo code of enhanced firefly optimization algorithm is detailed below.

3.3.1. Pseudo code of enhanced firefly optimization algorithm

Initialize firefly optimization algorithm parameters;
 MaxGen \rightarrow maximum number of generations.
 $t = 1$.
 Objective-function $\rightarrow f(x)$, where $x = (x_1, \dots, x_d)^T$.

Generate initial population of fireflies or $x_i(1,2, \dots, n)$.

Define light intensity l_i at x_i via $f(x_i)$ and distance between the fire-flies j and i is calculated using Chi square distance

While ($t < \text{MaxGen}$);

For $i = 1$ to n ;

For $j = 1$ to n ;

If ($l_j > l_i$), move firefly i towards j ; end if

Determine new solutions and update light intensity;

End for j ;

End for i ;

Rank the fire-flies and identify the current best;

End while;

Post process; visualization and result;

End procedure;

3.4 Classification

After selecting the optimal features, classification is accomplished by using DBN for classifying the facial images. The DBN is an important methodology among all deep learning models, which have a large number of Restricted Boltzmann Machines (RBMs). The learned activation unit of one RBM is utilized as the “data” for the succeeding RBM in the stack. Also, it is an undirected graphical approach in which visible variables x are linked to hidden unit y by using undirected weight. Where, the DBN are constrained, so there is no connection within the hidden or visible variables. The Eq. (18) shows a probability distribution over x , $y(p(x,y))$ and energy function is indicated as $E(x,y;\theta)$. The binary RBM is mathematically denoted in Eq. (18).

$$-\log p(x,y) \alpha E(x,y;\theta) = -\sum_{i=1}^{|P|} \sum_{j=1}^{|Q|} w_{ij} x_i y_j - \sum_{i=1}^{|P|} b_i x_i - \sum_{j=1}^{|Q|} a_j y_j \quad (18)$$

Where, $\theta = (w, b, a)$ is denoted as parameter set, w_{ij} is indicated as symmetric weight between the visible units, j is indicated as hidden unit, b_i and a_j are denoted as bias. In DBN, the number of visible and hidden layer is considered as $|P|$ and $|Q|$. The configuration n makes it easy to process the conditional probability distribution, where x and y are fixed that are mathematically described in the Eqs. (19) and (20).

$$p(y_j|x;\theta) = \text{sigm} \sum_{i=1}^{|P|} w_{ij} x_i + a_j \quad (19)$$

$$p(x_i|y;\theta) = \text{sigm} \sum_{j=1}^{|Q|} w_{ij} y_j + b_i \quad (20)$$

Where, $\text{sigm}(X) = \left(\frac{1}{1+e^{-x}}\right)$ is indicated as sigmoid function. The parameters $\theta = (w, b, a)$ is learned using contrastive divergence. The reason for using DBN classifier is that the parameters θ obtained by using RBM define both $p(x|y,\theta)$ and prior distribution $p(y|\theta)$. Therefore, the probability of developing visible variables are given in Eq. (21).

$$p(x) = \sum_y p(y|\theta) p(x|y,\theta) \quad (21)$$

After θ is learned from an RBM, $p(x|y,\theta)$ is kept. In addition, $p(y,\theta)$ is replaced by a consecutive RBM, which treats the hidden layer of the previous RBM as a visible value.

4. Result and discussion

In the experimental segment, the proposed work was simulated by utilizing MATLAB (version 2018a) with 3.0 GHZ-Intel i3 processor, 2TB hard disc, and 4 GB RAM. The proposed work performance was related with the existing works [13, 14] to estimate the effectiveness and efficiency of the proposed work. The performance of the proposed work was validated in light of accuracy, precision, and recall.

4.1 Performance metric

Performance metric is the procedure of collecting, reporting and analysing information about the performance of a group or individual. Mathematical equation of precision, accuracy, and recall are denoted in the Eqs. (22), (23), and (24).

$$\text{Precision} = \frac{TP}{FP+TP} \times 100 \quad (22)$$

$$\text{Accuracy} = \frac{TP+TN}{FN+TP+TN+FP} \times 100 \quad (23)$$

$$\text{Recall} = \frac{TP}{FN+TP} \times 100 \quad (24)$$

Where, TP is denoted as true positive, FP is represented as false positive, FN is indicated as false negative, and TN is denoted as true negative.



Figure.3 Sample images of ORL database

Table 1. Performance evaluation of proposed work with dissimilar performance measures on ORL database

Classifier	Precision (%)	Recall (%)	Accuracy (%)
SVM	61.2	67.09	70.98
KNN	72.39	78.01	79
LSTM	91.90	93.98	90.87
DBN	96	97.96	98.92

4.2 Quantitative analysis on ORL dataset

In this segment, ORL dataset is used to evaluate the performance of the proposed work. The ORL facial dataset comprises of 400 facial images with 40 individuals, each individual subject includes 10 face images. For some individuals, the facial images are taken at dissimilar times, lighting variations, facial expressions (smiling, not smiling, eye open, and eye closed) and facial details (glass/no glass). The sample images of ORL database is represented in Fig. 3.

In Table 1, the performance of the proposed work is evaluated by means of precision, accuracy, and recall on ORL database. Here, the performance evaluation is validated with 70% training of data and 30% testing of data. The validation outcome shows that the DBN classifier out-performed the existing classification methodologies like Support Vector Machine (SVM), K-Nearest neighbour (KNN) and Long Short Term Memory (LSTM). The precision of DBN classifier is 96% and the comparative classification methodologies: SVM, KNN and LSTM delivers 61.2%, 72.39 % and 91.90% of precision. Similarly, the recall of DBN classifier is 97.96% and the comparative classification methods delivers 67.09%, 78.01% and 93.98% of recall. Additionally, the accuracy of DBN is 98.92% and the comparative classifiers: SVM, KNN and LSTM attains 70.98%, 79% and 90.87% of accuracy. Graphical representation of proposed work by means of precision, recall, and accuracy on ORL dataset is indicated in Fig. 4.

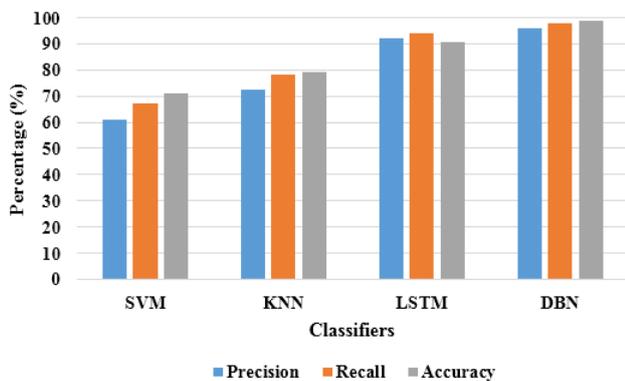


Figure.4 Graphical comparison of proposed work with dissimilar classifiers on ORL dataset



Figure.5 Sample images of YALE database

4.3 Quantitative analysis on YALE dataset

In this section, YALE dataset is used to evaluate the performance of the proposed work. The YALE database comprises of 165 grayscale images of 15 individuals. There are 11 images per subject, one per different facial expressions or configurations like center-light, happy, sad, sleepy, with/without glasses, normal, surprised, and wink. The sample images of YALE dataset are indicated in Fig. 5.

In Table 2, the performance of the proposed work is evaluated on YALE database. From the experimental investigation, the accuracy of DBN is 97.92%, and the existing classification approaches (SVM, KNN and LSTM) achieves 67.20%, 82.30% and 90%. Respectively, the precision, and recall of DBN classification approach is 98.90%, and 97%. In contrast, the comparative classifiers achieve minimum precision and recall compared to DBN. Graphical representation of proposed work by means of precision, recall, and accuracy on YALE dataset is indicated in Fig. 6.

Table 2. Performance evaluation of proposed work with dissimilar performance measures on YALE database

Classifier	Precision (%)	Recall (%)	Accuracy (%)
SVM	68.27	70.92	67.20
KNN	75.88	80.80	82.30
LSTM	92.93	94.98	90
DBN	98.90	97	97.92

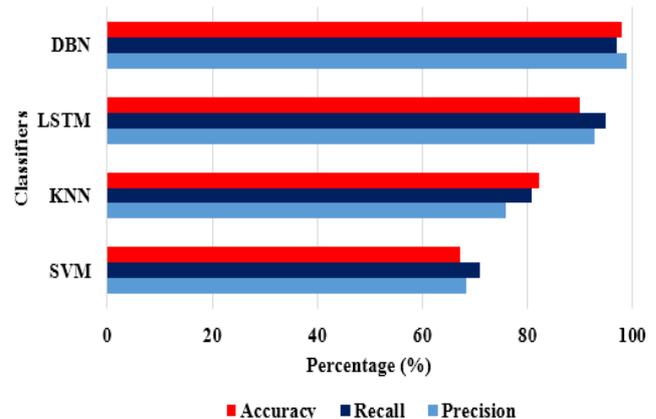


Figure.6 Graphical comparison of proposed work with dissimilar classifiers on YALE dataset



Figure.7 Sample facial images of FASSEG dataset

4.4 Quantitative analysis on FASSEG dataset

In this section, FASSEG dataset is used for evaluating the performance of the proposed work. The FASSEG dataset comprises of four subsets (multipose01, frontal03, frontal02, and frontal01), which includes several facial images. In face segmentation, FASSEG dataset is very useful for training and testing the automatic methods. The subset (multipose01) consists of 200 facial images with ground-truth masks for six classes (background, mouth, eyes, nose, skin, and hair). Consecutively, the subsets (frontal01 and frontal02) consists of 70 RGB images with labelled ground truth masks. At last, the frontal03 subset contains 150 face masks, which are captured in different face expressions, orientations, and illumination conditions. The sample collected facial images of FASSEG dataset is denoted in Fig. 7.

In Table 3, the proposed work performance is evaluated on FASSEG database. From the experimental study, the accuracy of DBN is 98.12%, and the existing classification approaches (SVM, KNN and LSTM) achieves 68.18%, 84.98% and 93.22%.

Similarly, the precision, and recall of DBN classifier is 94.09%, and 96.5%. In contrast, the comparative classifiers achieve minimum precision and recall compared to DBN. Graphical representation of proposed work by means of precision, recall, and accuracy on FASSEG database is specified in Fig. 8.

Table 3. Performance evaluation of proposed work with dissimilar performance measures on FASSEG database

Classifier	Precision (%)	Recall (%)	Accuracy (%)
SVM	80.83	87.93	68.18
KNN	77.09	88	84.98
LSTM	89	91.25	93.22
DBN	94.09	96.5	98.12

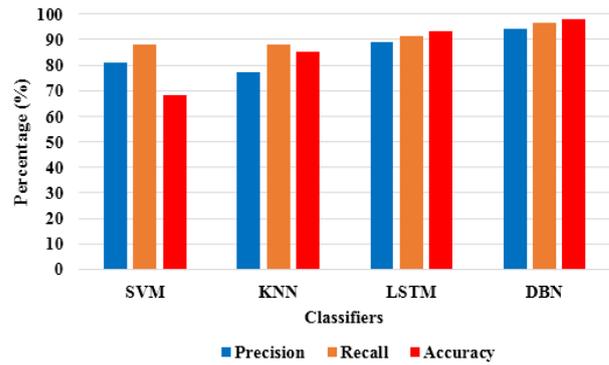


Figure.8 Graphical comparison of proposed work with dissimilar classifiers on FASSEG dataset

4.5 Comparative analysis

The comparative study of proposed and existing work is represented in the Table 4. L. Zhou, W. Li, Y. Du, B. Lei, and S. Liang [13] developed an Illumination Invariant Local Binary Descriptor Learning (IILBDL) method for face recognition. In this research, exponential discriminant analysis was utilized as a pre-processing method that enhances the discriminative ability of the mage by enlarging the margin between dissimilar classes relative to the same class. In the experimental phase, the developed method attained 91.48% of recognition accuracy in YALE dataset. In addition, A. Ouyang, Y. Liu, S. Pei, X. Peng, M. He, and Q. Wang [14] developed a hybrid approach for face recognition by combining PNNs and IKLDA. The developed method effectively (IKLDA+PNN) enhances the precision value and computing efficiency. In this work, the experimental investigation was conducted on the AR, ORL, and YALE datasets, which comprises of a wide variety of facial details, and expressions.

In the experimental consequence, the developed work averagely achieved 91.31% of accuracy in ORL dataset, and 77% of accuracy in YALE dataset. Compared to the existing work, the proposed work attained 98.92% of accuracy in ORL dataset, and 97.92% of accuracy in YALE dataset, which were superior related to the existing work. As discussed in the proposed work, feature optimization is a fundamental part of facial recognition in this article.

Table 4. Comparative study of proposed and existing work

Methodology	Datasets	Accuracy (%)
IILBDL [13]	YALE	91.48
IKLDA+PNN [14]	ORL	91.31
	YALE	77
Proposed work	ORL	98.92
	YALE	97.92

Whereas, several feature vectors are achieved under the feature extraction using BRISK and LTP, so enhanced firefly optimizer methodology is used for selecting the active features, which are fit for better classification.

5. Conclusion

The main aim of this experimental study is to propose a proper feature optimization and classification algorithm for recognising the facial images from the ORL, YALE, and FASSEG datasets. In this research, BRISK, and LTP features are utilized for extracting the features from the denoised facial images. Then, enhanced firefly optimization algorithm is used for selecting the optimal feature vectors. By selecting the optimal features from the extracted features, a set of most dominant discriminative features are obtained. These optimal features are classified by using DBN classifier. Compared to the existing work, the proposed work delivered an effective performance by means of quantitative and comparative analysis. From the experimental study, the proposed work attained 98.92% of accuracy in ORL dataset, 97.92% of accuracy in YALE dataset, and 98.12% of accuracy in FASSEG dataset, but the existing work attained only limited recognition accuracy. In the future work, a new hybrid optimization algorithm is developed for further improving the performance of face recognition.

References

- [1] M. Iqbal, M.S.I. Sameem, N. Naqvi, S. Kanwal, and Z. Ye, "A deep learning approach for face recognition based on angularly discriminative features", *Pattern Recognition Letters*, Vol. 128, pp. 414-419, 2019.
- [2] W. Yang, X. Zhang, and J. Li, "A Local Multiple Patterns Feature Descriptor for Face Recognition", *Neurocomputing*, 2019.
- [3] M. Kas, Y. Ruichek, and R. Messoussi, "Mixed neighborhood topology cross decoded patterns for image-based face recognition", *Expert Systems with Applications*, Vol. 114, pp. 119-142, 2018.
- [4] Y. Li, W. Zheng, Z. Cui, and T. Zhang, "Face recognition based on recurrent regression neural network", *Neurocomputing*, Vol. 297, pp. 50-58, 2018.
- [5] F. Cao, H. Hu, J. Lu, J. Zhao, Z. Zhou, and J. Wu, "Pose and illumination variable face recognition via sparse representation and illumination dictionary", *Knowledge-Based Systems*, Vol. 107, pp. 117-128, 2016.
- [6] B. Li, and G. Huo, "Face recognition using locality sensitive histograms of oriented gradients", *Optik-international journal for light and electron optics*, Vol. 127, No. 6, pp. 3489-3494, 2016.
- [7] W.H. Lin, P. Wang, and C.F. Tsai, "Face recognition using support vector model classifier for user authentication", *Electronic Commerce Research and Applications*, Vol. 18, pp. 71-82, 2016.
- [8] X. Wu, Q. Li, L. Xu, K. Chen, and L. Yao, "Multi-feature kernel discriminant dictionary learning for face recognition", *Pattern Recognition*, Vol. 66, pp. 404-411, 2017.
- [9] G. Feng, H. Li, J. Dong, and J. Zhang, "Face recognition based on Volterra kernels direct discriminant analysis and effective feature classification", *Information Sciences*, Vol. 441, pp. 187-197, 2018.
- [10] H. Gan, "A noise-robust semi-supervised dimensionality reduction method for face recognition", *Optik*, Vol. 157, pp. 858-865, 2018.
- [11] A. Alelaiwi, W. Abdul, M.S. Dewan, M. Migdadi, and G. Muhammad, "Steerable pyramid transform and local binary pattern based robust face recognition for e-health secured login", *Computers & Electrical Engineering*, Vol. 53, pp. 435-443, 2016.
- [12] W. Yang, Z. Wang, and B. Zhang, "Face recognition using adaptive local ternary patterns method", *Neurocomputing*, Vol. 213, pp. 183-190, 2016.
- [13] L. Zhou, W. Li, Y. Du, B. Lei, and S. Liang, "Adaptive illumination-invariant face recognition via local nonlinear multi-layer contrast feature", *Journal of Visual Communication and Image Representation*, Vol. 64, pp. 102641, 2019.
- [14] A. Ouyang, Y. Liu, S. Pei, X. Peng, M. He, and Q. Wang, "A hybrid improved kernel LDA and PNN algorithm for efficient face recognition", *Neurocomputing*, 2019.
- [15] Q. Liu, C. Wang, and X.Y. Jing, "Dual multi-kernel discriminant analysis for color face recognition", *Optik*, Vol. 139, pp. 185-201, 2017.
- [16] C.H. Yoo, S.W. Kim, J.Y. Jung, and S.J. Ko, "High-dimensional feature extraction using bit-plane decomposition of local binary patterns for robust face recognition", *Journal of Visual Communication and Image Representation*, Vol. 45, pp. 11-19, 2017.
- [17] I.M. Revina, and W.S. Emmanuel, "Face Expression Recognition with the Optimization based Multi-SVNN Classifier and the Modified LDP Features", *Journal of Visual*

Communication and Image Representation, Vol. 62, pp. 43-55, 2019.

- [18] H. S. Suresha and S. S. Parthasarathi, "ReliefF feature selection based alzheimer disease classification using hybrid features and support vector machine in magnetic resonance imaging", *International Journal of Computer Engineering & Technology (IJCET)*, Vol. 10, No. 1, pp. 124-137, 2019.
- [19] S. Leutenegger, M. Chli, and R. Siegwart, "BRISK: Binary robust invariant scalable key points", In: *Proc. of IEEE International Conference on Computer Vision (ICCV)*, pp. 2548-2555, 2011.
- [20] V. Naghashi, "Co-occurrence of adjacent sparse local ternary patterns: A feature descriptor for texture and face image retrieval", *Optik*, Vol. 157, pp.877-889, 2018.
- [21] ORL dataset:
<http://www.cad.zju.edu.cn/home/dengcai/Data/FaceData.html>
- [22] YALE dataset:
<https://www.kaggle.com/olgabelitskaya/yale-face-database>
- [23] FASSEG dataset:
<http://massimomauro.github.io/FASSEG-repository/>