Check for
updates

Improved crossover firefly algorithm based deep Beleif network for low-resolution face recognition

Wael Mohammad Alenazy [1] · Abdullah Saleh Alqahtani [1]

Abstract

Face detection by low-resolution image (LR) is one of the key aspects of Human-Computer Interaction(HCI). Due to the LR image, which has changes in pose, lighting, and illumination, the performance of face recognition is reduced. In this work, we propose the Deep Belief Network-Crossover based Firefly (DBN-CROFF) method for face recognition from low-resolution images. The Histogram of Gradient (HOG) and 2-Dimensional Discrete Wavelet Transform (2D-DWT) to extract facial width, size of the cheeks, skin tone, nose, and lip shape features from facial data. The Kernel Principle Component Analysis (k-PCA) is used to successfully reduce the dimension of the feature. The experimental performance of the proposed method is evaluated using four datasets namely LFW, Multi-PIE, Extended Yale-B, and FERET with conventional techniques. Finally, the proposed DBN-CROFF solution surpasses the other conventional facial recognition approaches by giving a higher accuracy of recognization.

Keywords Low-resolution image · Face recognition · Facial features · DBN · CROFF

1 Introduction

One of the important research fields in computer vision is face recognition that made many valuable contributions in recent years. Face recognition systems are used to recognize a person from a mug, video, shot, and digital images. Recently, most commonly and popularly used biometric technology is face image recognition. Generally, the face recognition model mainly concentrates on large sufficient frontal faces that consist of enough data for recognition [11]. The attributes of face recognition technologies are significantly higher reliability and low invasiveness of the acquisition. One of the major challenges faced by the facial recognition systems is the identification of faces from the LR images [40] captured from a longer distance. The discriminate face properties are tainted due to the low-resolution image that considerably reduces the traditional face recognition accuracy.

✉ Wael Mohammad Alenazy
   walenazy@ksu.edu.sa; wael.alenazym@gmail.com

[1]   Department of Self Development Skills, CFY Deanship King Saud University, Riyadh, Saudi Arabia

Practically, the face recognition model deals with poor image quality or low resolution (LR) images and also larger variations in lighting conditions, facial expressions, and pose. Hence, the face recognition using low resolution is a challenging task. The conservative face recognition models are working well for high-quality images under suitable environmental conditions [1]. By using deep learning methods, the high-resolution face images produce over 99% of recognition rates. Nevertheless, face images including limited information, which for Low-Resolution Face Recognition(LR-FR) [26, 39] are different from others. If the face resolution is poor, the identification rate decreases significantly. The challenging issues in face recognition are due to face misalignment, noise effects, and lack of resolution robust features. Hence, it is a challenging research area in pattern recognition and computer vision [2, 5].

Real-time and high reliability with automatic face recognition is difficult in the computer vision community. The deep learning models have been widely used to enhance low-resolution face recognition performance in past years. LR image features are optimized using different methods and these methods adopt different parameters to optimize the LR images [21]. Most of the works are focused on LR image transformation into a high resolution (HR) image that is applied to enhance the quality of the LR image. The photos with super-resolution are used as a way of reconstruction for the HR images of the particular LR test image and the identification of the face. Although, face recognition methods efficiently deal with machine learning algorithms such as Deep Convolutional Neural Network (DCNN), Convolutional Neural Network (CNN), and Deep Learning (DL) algorithms [40]. In recent years, evoluationary methods has been achieved better performance in enginnering applications [3, 4, 25, 16, 17, 28, 29, 32, 33, 30]. In this paper, we suggested the Deep Belief Network-Crossover based Firefly (DBN-CROFF) method for LR-FR. The major contribution of this research is summarized as shown below:

- The exploitation capacity of Firefly Algorithm (FA) is improved with the usage of crossover operation and the newly developed algorithm is named as Crossover based Firefly (CROFF) algorithm.
- The Histogram of Gradient (HOG) and 2-Dimensional Discrete Wavelet Transform (2D-DWT) to extract facial width, size of the cheeks, skin tone, nose, and lip shape features from facial data.
- The kernel- Principle Component Analysis (k-PCA) is utilized for dimensionality reduction. We used LFW, Multi-PIE, Extended Yale-B, and FERET dataset for experimental analysis.
- The proposed DBN-CROFF method with FERET, Extended Yale-B, Multi-PIE, and LFW dataset provided 93%, 95%, 90%, and 89% accuracy rates

The remainder of this paper is arranged in the following structure where the literature review is specified in Section 2 and section 3 formulates the Crossover based Firefly (CROFF) algorithm. The proposed DBN-CROFF method for face recognition is explained in section 4. Section 5 addresses the results and Section 6 concludes the paper.

2 Review of related works

A novel approach for LR-FR was presnted by Shakeel Saad et al. [26]. Initially, the sparse error matrix and low-rank matrix are decomposed. By using the sparse coding based algorithm

to learn the projection matrix. The similarity between the query image and projected gallery image features are determined with the usage of the coefficient vector depending upon the linear regression. The novel pre-processing model such as a morphological model was developed to enhance the visual image quality. Experimental results demonstrated a better and accurate LR face recognition model. Zangeneh et al. [39] proposed a deep convolutional neural network (DCNN) based novel coupled mappings to recognize the face from an LR image. By using non-linear transformation, the DCNN is divided into two sections that map low and high-resolution images into the common space. A total of 14 layers are used corresponding to the LR image transformation. The network training is performed by the backpropagation of low and high images correspond to the distance between the features. The recognition performance is improved in terms of visual quality than other methods.

A multi-scale patch-based feature learning (MSPRFL) model was developed by Goa et al. [8], which can solve the LR face reorganization problems. Using training datasets, the MSPRFL model generates multi-stage data to a specific, reliable feature resolution model that minimizes discrepancy in resolution. Several common face data sets are used to improve the MSPRFL model's efficiency in LR facial recognition. Marciniak et al. [22] used a novel method to identify a face from an LR image and also the reliability of the real-time system is analyzed. This research results in the identification of the face and the extraction of the facial features in the image. The influence of LR-FR provided higher recognition performance. Discriminant Correlation Analysis (DCA) for LR face recognition was suggested by Haghighat et al. [10]. The feature correlation in high and LR images is analyzed by using Discriminant Correlation Analysis. At the same time, the DCA method maximizes the correlation between both feature sets. The DCA is effective in real-time application and offers better performance in face recognition.

Ullah et al. [31] proposed a new method to identify LR and occlusive facial images almost in real-time. The facial images that appear in multimedia applications typically have the following characteristics, such as posture, occlusion, expression, LR, and uniform illumination. There are 68 points to trace the input face image for face recognition operation. Linear Discriminant Analysis and adaptive booting models are generated for feature extraction and the face feature classifications are performed via classic nearest center classifier. A sharped classifier deep neural network with multi hierarchy loss was presented by Sun et al. [27]. They also suggested a shared classifier between LR and HR face images. Multi hierarchical loss is integrated into intermediate layers and intermediate feature overuse is used to avoid reductions in accuracy and also to enhance face recognition performance.

3 Formulation of crossover based firefly algorithm (CROFF)

In this section, we formulate the Firefly algorithm and Crossover operation namely Crossover based Firefly Algorithm (CROFF).

3.1 Firefly optimization algorithm (FA)

The light emission ability of fireflies is inspired by Firefly Algorithm (FA) [34]. The different goals such as food search and reproduction are achieved by initiating inter-firefly communication. In order to solve the optimization problems, the light intensity produced by fireflies with the objective function (fitness), is said to be optimized [23]. The light intensity emitted by

the firefly $I(d)$ based on the physics rules of lights. The fireflies at distance d are determined using Eq. (1).

$$I(d) = \frac{I_0}{d^2} \tag{1}$$

From the above equation, the light intensity created at the light source is represented as I_0. At distance d, the light intensity I and the absorption coefficient of the medium are χ as explained in Eq. (2).

$$I = I_0 \exp(-\chi d^2) \tag{2}$$

Here, the distance between the light observation pint and the light source is denoted as χ. Light intensity is correlated with the attraction of the fireflies in FA.

$$A = \alpha_0 \exp(-\chi dn)(n >= 1) \tag{3}$$

The attraction value located at a zero distance is taken as α_0 and as shown in the following equation, the Euclidian distance between two butterflies(a_i and b_j) is estimated as follows

$$d_{ij} = |a_i - b_j| = \sqrt{\sum_{r}^{k=1} (a_{i,k} - b_{j,k})^2} \tag{4}$$

The firefly position is changed in each generation based on Eq. (5).

$$Y_i = a_i + \alpha_0 \exp\left(-\chi d_{ij}^2\right)(a_i - b_j) + \beta\varepsilon \tag{5}$$

Therefore, the randomization parameter is χ and ε is the random number obtained using the Gaussian distribution. The randomization parameter is used to control the solution search space.

3.2 Crossover based firefly algorithm (CROFF)

The exploitation capacity of Firefly Algorithm (FA) is improved with the usage of crossover operation and the newly developed algorithm is named as Crossover based Firefly Algorithm (CROFF). The "DE/Best/1" mutation model there in the Differential Evolution (DE) is utilized by crossover. The parallel direct search model in DE utilizes the NP-D dimensional parameter vector [38].

$$b_j, I \quad, j = 1, 2, \ldots\ldots, NP \tag{6}$$

According to Eq. (6), the total number of generations used on NP is I and the target vector b_j. Hence, the NP value becomes constant, and taking the weighted difference among two population vectors in DE creates a new parameter (i.e. mutation). The parameters of mutated vectors are added with another predetermined vector parameter to form a trial vector. The crossover is the mixing of these parameters. The global optimal solution of the reference point is set by using "DE/best/1" then select two individual solutions such as $A_{d1}^t = \left[B_{d1,1}^t, B_{d1,2}^t, \ldots\ldots\ldots, B_{d1,k}^t, \ldots\ldots\ldots, B_{d2.D}^t\right]$ and $A_{d2}^t = \left[B_{d2,1}^t, B_{d2,2}^t, \ldots\ldots\ldots, B_{d2,k}^t, \ldots\ldots\ldots, B_{d2.D}^t\right]$ to obtain the weight difference that is represented in Eq. (7).

$$V_{d,k}^t = gaA_d + rand(0,1) \times \left(A_{d1,k}^t, A_{d2,k}^t\right) \qquad (7)$$

Hence, the mutant solution is $d = 1,2,....,$ $NP,$ $k = 1,2,...,$ $D,$ $V_{d,1}^t = \left[v_{d,1}^t, v_{d,2}^t, ..., v_{d,k}^t, ..., v_{d,D}^t\right]$. The global optimal solution of dth component is gaA_d. Thereafter, calculate the new velocity of each solution in the evolution procedure. The novel crossover operation is executed as CROFF that combines the promising data. Therefore, the crossover operation utilized in the proposed CROFF algorithm is explained in Eq. (8).

$$B_{d,k}^{t+1} = \begin{cases} B_{d_3,k}^t + V_{d,k}^{t+1} & , if\ rand(0,1) < CR \\ gaB_d + rand(0,1) \times \left(B_{d1,k}^t, B_{d2,k}^t\right) & otherwise \end{cases} \qquad (8)$$

Here, the crossover rate is CR that manages the inheriting probability from the new location. In the final result, the CROFF has inherited some components from the new position's current solution. Some parameter utilizes the "DE/Best/1" mutation model for the probability of (1-CR) with a new solution. The CROFF can accomplish optimal exploitation ability as the "DE/Best/1" mutation operator denotes good exploitation ability. The CROFF algorithm pseudocode is formulated in Algorithm 1.

Algorithm 1: CROFF algorithm

t=0
Fitness value=0
Initialization of a random population of fireflies and its parameters
while
 Fitness value < Max-Fitness **do**
for
j=1 to NP **do**
      Calculate the light intensity *I(d)* of fireflies based on equation (1)
      Calculate the attractiveness between the fireflies using equation (3)
      Compute the Euclidean distance between two fireflies using equation (4)
      Estimate the crossover operation based on equation (8)
      Estimate the dth solution
if
    A_d^t is better than **then**
    $gaA = A_d^t$
end if
   Fitness value= Fitness value+1
end for
      t=t+1
end while

---

4 Proposed DBN-CROFF algorithm for LR-FR

The proposed DBN-CROFF algorithm for LR image facial recognition, as seen in Fig. 1 is used to identify the faces present in the LR image. Before passing data into DBN, the proposed work utilizes various feature extraction and dimensionality reduction process. The facial width,
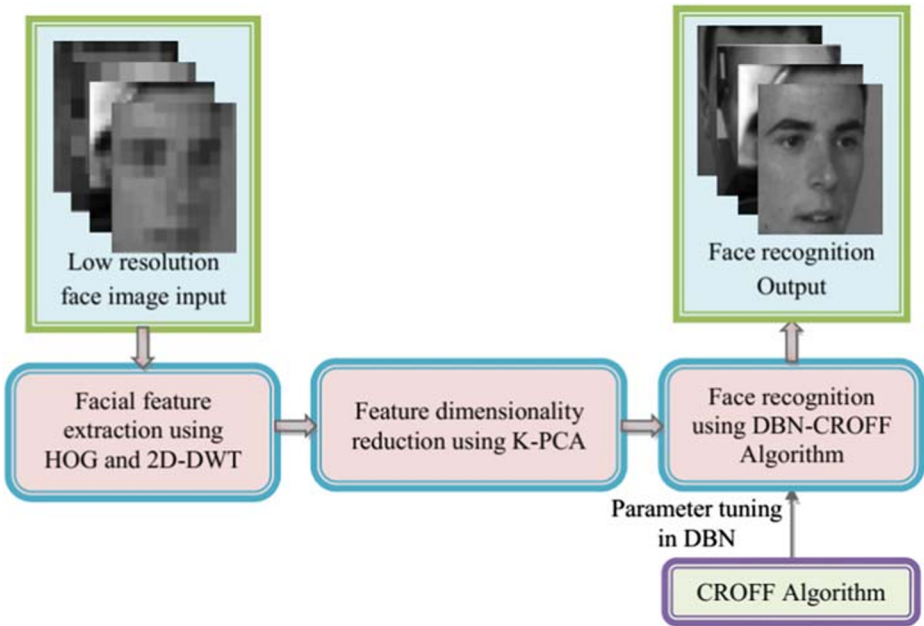
Fig. 1 Proposed DBN-CROFF algorithm structure for face recognition

size of the cheeks, skin tone, nose, and lip shapes from facial data based on the characteristics of different faces are obtained for data acquisition. Histogram of Gradient (HOG) and 2-Dimensional Discrete Wavelet Transform (2D-DWT) is the feature extraction methods used to retrieve the facial features from the LR image. The information regarding the image gradients is extracted using HOG to formulate the structure shape in an image [36].

In the year of 2005, HOG was initially used for pedestrian detection. HOG splits the images into $M \times M$ overlapped cell pixels and the cells are arranged like overlapping blocks. Horizontal and vertical kernels such as $[-1, 0, 1]^T$ and $[-1, 0, 1]$ are used to calculate the pixel gradients inside each cell. The varying cell size allows the processing of information at various scales. The orientation details are recorded by using the number of orientation histogram bins. Histogram to compress the descriptors of the feature arranges a vector of gradients that generalizes information about the cells. HOG is invariant to photometric and geometric transformations by normalizing the gradient vectors. The local shape by capturing the gradient structure and edge are characterized due to HOG and the features are smaller contrasted to orientation or local spatial bin size that is invariant to rotations and translations.

The image is broken down into a series of simple functions (wavelets) using the 2D-DWT, which provides information on time and frequency [7]. The set of filter banks implements the 2D wavelet decomposition of an image. Initially, apply the 1D-DWT along with the columns and rows. The cascaded design of low pass filter (LPF) and high pass filter (HPF) is located in the filter banks. The four sub-band images such as High High (HH), High Low (HL), Low High (LH), and Low Low (LL) are the decomposition results. In this work, we utilize ten different feature vectors including facial width, cheeks size, skin tone, nose shape, and lip shape to calculate the relevant feature extraction method. This feature vector dimensionality is again reduced to accelerate the DBN training then the minimized features are applied to the input of semi-supervised DBN and the recognition rate for different feature vectors is

computed. Reduce the HOG and 2D-DWT features of facial width, cheeks size, skin tone, nose shape, and lip shape that assign ranking based on the accuracy. The proposed DBN-algorithm is used to predict the most appropriate facial features from LR images with low computational complexity.

4.1 Kernel principal component analysis (K-PCA) for dimensionality reduction

The extraction of features and the design processes are generally carried out by reducing dimensionality. The problem of dimensionality is induced by the size of the data and increased computational complexity. Therefore, owing to the burden of dimensionality the classification process is complicated and hence reduction in dimension is required. We used k-PCA to reduce the dimensionality of these features in this work. Initially, all the input points P is mapped into $F(P)$ in the huge dimensional feature space. The covariance matrix is explained in Eq. (9).

$$\sum F = \frac{1}{M} \sum_{m=1}^{M} F(P_m) F(P_m)^T \tag{9}$$

For the covariance matrix, substitute the above equation into the Eigen equation.

$$\sum_F \phi_i = \chi_i \phi_{ji} \tag{10}$$

Hence, the eigenvector is ϕ and the linear combination of M mapped data points obtains the eigenvector.

$$\phi_i = \sum_{m=1}^{M} \delta_m^{(i)} F(P_m) \tag{11}$$

$$\delta_m^{(i)} = \frac{1}{\chi_i M} (F(P_m \cdot \phi_i)) \tag{12}$$

By multiplying $F(P_m)$ in the left side of Eq. (11), which is formulated as below:

$$(F(P_n) \bullet \phi_i) = \sum_{m=1}^{M} \chi_m^{(i)} k(P_n, P_{nm}) \tag{13}$$

From the above equation, the kernel model of internal products of two vectors determined on the dimensional space F is $k(P_n, P_m) = (F(P_n). F(P_m))$, where $m, n = 1, . , M$. When the value of $n = 1, ..., M$ means the Eq. (13) is the nth component.

$$\chi_i M \chi_i = k \chi_i \tag{14}$$

Hence, the $M \times M$ matrix is taken as k that comprises of $k(P_j, P_i)$ kernel elements. From this, $\chi_i = \left[\chi_1^{(i)}, \ldots \ldots \ldots, \chi_M^{(i)} \right]^T$, where $i = 1, \ldots \ldots \ldots, M$. The eigenproblem of k is solved by using eigenvector M.

$$(F(P) \bullet \phi_i) = \sum_{m=1}^{M} \chi_M^{(i)} k(P_n, P_m) \tag{15}$$

The image dimensionality is reduced using kernel PCA (k-PCA) that sorts the points based on the optimal face feature class. Furthermore, this k-PCA model minimizes the memory consumption and computational time.

4.2 Restricted Boltzmann machine (RBM)

One of the important types of generative stochastic Artificial Neural Network (ANN) us Restricted Boltzman Machine (RBM) that is used to fit the input data probability distribution. RBM consists of two layers namely the input layer (visible layer) and the hidden layer. The neurons present in the input layer has a part of input in it. In the input layer, there are some neurons with input data. The hidden layer [12] also comprises certain neurons that are identical to the visible layer output. While the neurons in the same layer are limited, the neurons of the Hidden Layer(S) and the Visible Layer(L) are connected by certain edges. The energy function completes the probability distribution between the hidden and the input layers that are known by the RBM. Therefore, the energy function is as shown in Eq. (16).

$$
\begin{aligned}
E(L,S) &= -S^T W L - c^T L - d^T S \\
&= -\sum_j c_j L_j - \sum_i d_i S_i - \sum_{ji} W_{ji} L_j S_i
\end{aligned}
\tag{16}
$$

Hence, j^{th} visible layer is L_j and the ith input layer is S_i. The weight connection between the L_j and S_i is W_{ji}. The c_j and d_i values denote the unit threshold values used for both the hidden and the visible layer. The below equation explains the joint probability distribution.

$$
P(L,H) = \exp\left(-E(L,H)\right) / \left\{ \sum_{L,H} \exp(-E(L,H)) \right\}
\tag{17}
$$

The binary units are used in most of the studies. Hence, L_j and $S_i \in (0,1)$. The initial j hidden layer nodes probability is represented in Eq. (18).

$$
P(S_i = 1|L) = sigmoid\left(d_i + \sum_i W_{ji} S_i \right)
\tag{18}
$$

The probability function of the entire hidden layer is shown below:

$$
P(S|L) = \prod_i (S_i|L)
\tag{19}
$$

The activation probability of nodes in the visible layer and the entire visible layers are as shown in Eqs. (20) and (21).

$$
P(L_j = 1|S_i) = sigmoid\left(d_j + \sum_j W_{ji} L_j \right)
\tag{20}
$$

$$
P(L|S) = \prod_j (L_j|S)
\tag{21}
$$

The input vector probability L is the likelihood $P(L)$ over the hidden units. Therefore, Eq. (22) is set.

$$P(L) = \sum_S \left\{ \exp(-E(L,S)) / \sum_{L,S} \exp\left(-E\left(L,S\right)\right) \right\} \tag{22}$$

The Deep Belief Network (DBN) objective function is shown below:

$$L(\theta, L) = \sum_{L \in D} \log P(L, \theta) \tag{23}$$

Hence, the dataset to train the model is represented as $\theta \in \{c, d, W\}$;D. By using Bayesian statistic theory, the objective function shown in Eq. (23) is used for optimization. Maximize the log-likelihood to fit the model parameters that solve the optimization problem. By applying the principle of gradient descent to the objective function, the subsequent algorithms are derived.

$$(\partial \log P(L))(\partial w_{ji}) = E_P\left[S_i L_j\right] - E_{\widehat{P}}\left[S_i L_j\right] \tag{24}$$

$$(\partial \log P(L))/(\partial c_j) = E_P\left[L_j\right] - E_{\widehat{P}}\left[L_j\right] \tag{25}$$

$$(\partial \log P(v))/(\partial d_i) = E_P[S_i] - E_{\widehat{P}}[S_i] \tag{26}$$

The probability expectations over L under the model distribution $\hat{P}$ and the empirical distribution P are $E_{\hat{P}}[\cdot]$ and $E_P[\cdot]$. The initial terms of Eqs. (24)–(26) are calculated with the usage of Eq. (18) and (20) for a given RBM. Due to the expectation under the distribution $\hat{P}$, the second terms of Eqs. (24)–(26) are cannot be obtained directly. These expectations are calculated by the implementation of alternating Gibbs sampling. It is not applicable in practice because this is a time-consuming method. Therefore, the Hinton contrast divergence (CD) and fast learning algorithm were proposed to overcome this issue and identify the way out. Two tricks are used to improve the sampling process: (i) Markov chain initialization is finished by using training sample, (ii) the samples are got from the k-steps of Gibbs sampling (*i.e CD-k*). When $k = 1$, several existing methods have shown the satisfactory performance of the CD. From this, we implement CD-1 then update the parameters of $c, d,$ and W in the RBM as shown in Eq. (27)–(29).

$$W^{tm+1} = W^{tm} + \gamma\left(P\left(S|L^{(0)}\right)\left[L^{(0)}\right]^T - P\left(S|L^{(1)}\right)\left[L^{(1)}\right]^T\right) \tag{27}$$

$$c^{tm+1} = c^{tm} + \gamma\left(L^{(0)} - L^{(1)}\right) \tag{28}$$

$$d^{tm+1} = d^{tm} + \gamma\left(P\left(S|L^{(0)}\right) - P\left(S|L^{(1)}\right)\right) \tag{29}$$

Therefore, the learning rate and the time steps are represented as γ and *tm*. Algorithm 2 explains the pseudocode to train the RBM.

Algorithm 2: RBM training pseudocode

Input: Training sample $L^{(0)}$ from the distribution of training over the Restricted Boltzmann Machine.

Output: Parameters updated in RBM such as *c, d,* and *W*

for

   Each hidden units *i* **do**;

   Calculate $P(S_i^{(0)}1|L^{(0)})$ with Equation (18)

   Sample $S_i^{(0)} \in 0,1$ from $P(S_i^{(0)}1|L^{(0)})$

 end for

 for

   Each visible units *j* **do**;

   Calculate $P(L_j^{(1)}1|S^{(0)})$ with Equation (20)

   Sample $L_j^{(1)} \in 0,1$ from $P(L_j^{(1)}1|S^{(0)})$

end for

for

   Each hidden units *i* **do**;

   Calculate $P(S_i^{(1)}1|L^{(1)})$ by means of Equation (18)

   Sample $S_i^{(0)} \in 0,1$ from $P(S_i^{(1)}1|L^{(1)})$

 end for

   $W = W + \gamma(P(S|L^{(0)})[L^{(0)}]^T - P(S|L^{(1)})[L^{(1)}]^T$

   $c = c + \gamma(L^{(0)} - L^{(1)})$

   $d = d + \gamma(P(S|L^{(0)}) - P(S|L^{(1)}))$

return *c, d, W*

4.3 Proposed DBN-CROFF algorithm for face recognition

The unsupervised model of DBN [20] is composed of *m* layers. The initial layer is the visible or input layer with *c* input vectors and the final layer is the output layer O^l. The layers between the hidden layers are denoted by $O^1, \ldots, O^{l-1}$. The RBMs with the sigmoid function is utilized to represent the DBN hidden layers. The top activation functions must utilize the regression function to predict face features. The joint probability distribution over the hidden units and input vector *y* are learned using Eq. (30).

$$P\left(c, O^1, \ldots \ldots, O^n\right) = \left(\prod_{i=1}^{l-1} P\left(O^{l-1}|O^j\right)P\left(O^{l-1}, O^l\right)\right) \tag{30}$$

Hence, the process of recognition is important and the top layer features are distracted from the down layer features. These features differ less than raw face image data that is ideal for inference and informative purposes. Generally, the neural network with the multiple layers is learned with the help of backpropagation (BP) but it traps into local minima [14]. The DBN is

deeper due to the larger number of parameters and also the optimization problem is more difficult. Initialize few good values to the parameters by one possible method to alleviate the local minima backpropagation scheme. Where the target output $H_k \in R^D$ for an input training set (c_i, d_i), $1 \le i \le l$. The cost function of the training targets is minimized by taking the average negative log-likelihood.

$$
\begin{aligned}
J(W) &= -\log\left(\prod_{l,k} P\big(O_{l,k}|A, W\big)(P(W))\right) \\
&= -\sum_{l,k} \log P\big(O_{l,k}|A, W\big) - \log P(W)
\end{aligned}
\tag{31}
$$

Therefore, the A matrix comprises of data a_l and each linear parameters of DBN present in W. Thus, the covariance with the zero mean gaussian function is explained as below:

$$
J(W) = \frac{1}{2\sigma^2} \sum_{l,k} \big\|H_{l,k} - O_l(c_l)\big\|^2 + W^T \Sigma_p^{-1} W
\tag{32}
$$

The optimization function of multinoulli distribution is denoted in Eq. (33).

$$
J(W) = \frac{1}{2\sigma^2} \sum_{l,k} O_{l,k} \log(H_l(c_l)) + W^T W
\tag{33}
$$

For input c_l, the $H_l(c_l)$ represents the kth output of DBN. The parameter tuning in DBN is more important to increase the learning speed. We suggested the CROFF algorithm in this paper to refine the DBN parameters and it eliminates the inactivity problem during the DBN training process. The flow chart of the CROFF algorithm for DBN parameter optimization is demonstrated in Fig. 2.

The proposed DBN-CROFF algorithm for face recognition is explained as follows:

i.   Build the training dataset from the input data.
ii. Train the DBN according to the DBN technique mentioned in the section above.
iii. Initialization of CROFF parameters (Fireflies and crossover parameters).
iv. DBN train using different fireflies or CROFF parameters.
v.   Evaluate the fitness function of DBN as shown in Eq. (23).
vi. Choose the best firefly and update its position.
vii. Find if the termination condition met or not.
viii. If the condition met then stop the CROFF process and obtain the optimal parameters otherwise back to step (iv).

The face recognition is influenced by the LR image with poor quality. Hence, the features of facial width, cheeks size, skin tone, nose, and lip shape are taken from the facial dataset. The time-series autocorrelation function is implemented to identify the face. From the raw subset, the autocorrelation function is more helpful to select the informative feature subset and it constructs the training dataset [19]. Assume that, the raw data series (reduced feature dimensionality) are denoted as Y, where $Y = \{Y(tm) : tm \in T\}$. The index set is T and R_k is the autocorrelation coefficient as shown in Eq. (34).

$$
R_k = R(Y(tm), Y(tm-k)) = \frac{\sum_{tm=k+1}^{m} \left(Y(tm) - \overline{Y}\right)\left(Y(tm-k) - \overline{Y}\right)}{\sum_{tm=1}^{m} \left(Y(tm) - \overline{Y}\right)}
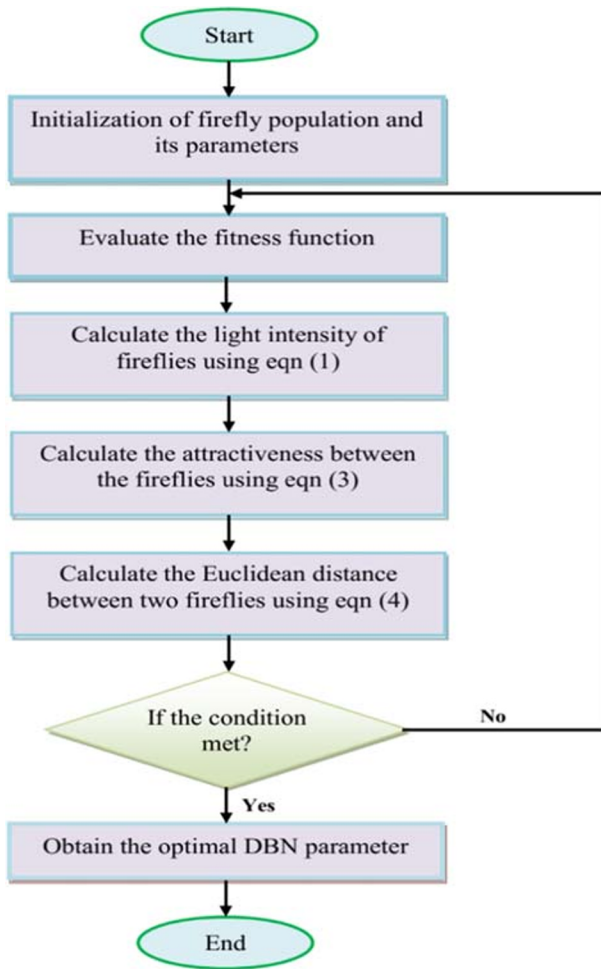\tag{34}
$$

Fig. 2 CROFF algorithm flowchart for DBN parameter optimization

Hence, the mean of Y is $\overline{Y}$. The linear correlation of timestamp tm and $tm\text{-}1$ are measured by R_k. The topmost features are selected as the candidate variables for face recognition. The LR images have few spatial dependencies so considering the correlation between these LR images. Therefore, for identifying the face, the top most features from the LR images are picked.

5 Result and discussion

This section evaluates the efficiency of the proposed DBN-CROFF algorithm for face recognition from LR images. The proposed research work is implemented using MATLAB software. The efficiency of the proposed DBN-CROFF algorithm is assessed by a number of experimental analyses which is discussed in the following section.

5.1 Database description

Low resolution face images are collected from LFW [14], Multi-PIE [9], Extended Yale-B [35] and FERET [24]. Therefore, each dataset comprises the image with variations in illumination, lighting conditions, facial expressions, and pose. In the initial stages of pre-processing the MTCNN will be used for detecting and aligning face images [37]. The HR images were mainly taken and downsampled to obtain the LR images of a specific size in this work. The sample image dataset for LR-FR is as shown in Fig. 3. Each dataset performance is described in the following section.

Multi-PIE dataset, (c) Extended Yale-B and (d) FERET dataset.

5.1.1 Performance of LFW dataset

The output of the proposed DBN-CROFF with the LFW dataset is evaluated in this section and facial images in the low resolution are taken in unregulated environments. Fig. 4 demonstrates the proposed DBN-CROFF performance for face recognition based on the LFW data set. Each image from the LFW dataset is obtained with different variations in lighting conditions, make-up, face, and expression. This dataset comprises of 12,469 face images from 47,103 individuals. From this, more than two images present in the 1680 individuals, and also 512 of them

Dataset name	*LR image*	*FR output*
LFW dataset		
Multi-PIE dataset		
Extended Yale-B		
FERET dataset		

Fig. 3 Sample dataset of face recognition from the low-resolution image, (**a**) LFW dataset, (**b**)
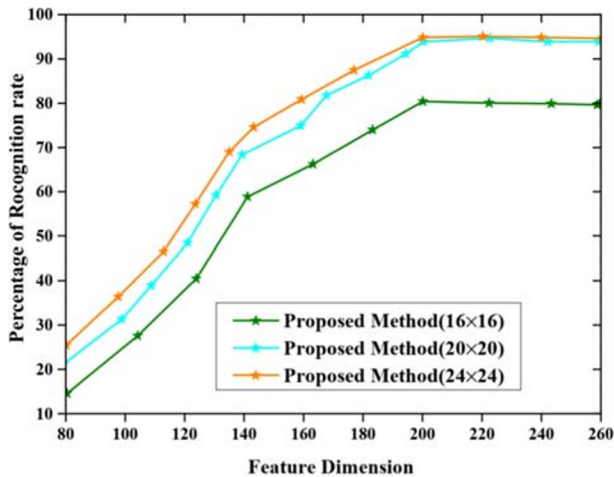
Fig. 4 Proposed DBN-CROFF face recognition performance based on the LFW dataset

comprise more than 4 images. We randomly select 120 subjects with 4 images that are collected from the CASIA-Web face dataset for the training process [18]. Provide the two images that are collected from the LFW dataset to construct the gallery set for the testing process, and the remaining two are used for the probe set. Note that, both training and testing datasets are different. The downsampling operation is performed using the low-resolution probe images of 16×16, 20×20 and 24×24 resolutions. Hence, the proposed DBN-CROFF algorithm outperforms 82.23% accuracy.

5.1.2 Performance of multi-PIE(M-PIE) dataset

Figure 5 explains the proposed DBN-CROFF face recognition performance based on the M-PIE dataset. The M-PIE dataset contains over 750,000 face images of 309 subjects and the face images were taken from different sessions. The experiment is carried out on the 04 session subset. The images comprise 20 illumination conditions under the frontal pose. In this experiment, a 05–1 camera with 01 recording numbers is used. For the training process, choose 50 subjects at random and the remaining subjects for the testing process. To construct the gallery set, each subject with six images is selected, and the remaining 14 images are included in the probe set. The downsampling technique is used to produce an image size 8×8 of an LR sample. The HR images of the M-PIE sample data set corresponding to the LR images are demonstrated in Fig. 3. The LR image recognition rate at various feature dimensions (M-PIE dataset) is formulated in Fig. 5. The proposed DBN-CROFF algorithm delivers better accuracy in recognition than other methods.

5.1.3 Performance of extended Yale-B(EYB)

Figure 6 shows the proposed performance of DBN-CROFF face recognition in the EYB data set. This EYB dataset comprises 2203 photographs of 39 subjects. Every subject consists of 65 photographs taken at different illumination conditions. For the training process, select 15, 20, and 25 per subject at random. The downsampling operation generates a low-resolution probe image size of 12×12. Fig. 6 shows that the increasing trend and the increasing feature
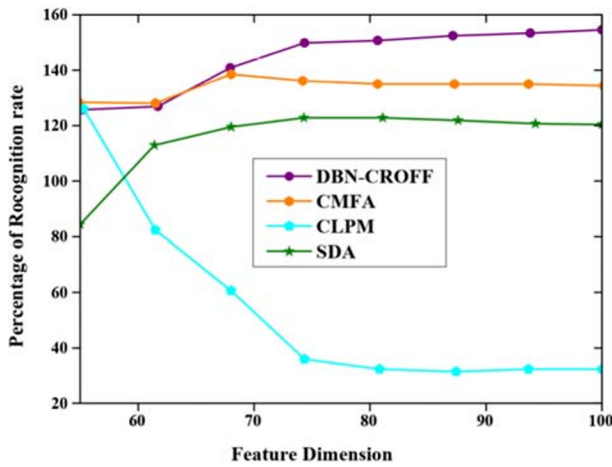
Fig. 5 Proposed DBN-CROFF face recognition performance based on the Multi-PIE dataset

dimension remains stable. The proposed DBN-CROFF method is compared with existing CMDA [40], CLPM [18], CDMMA [15], CCA [35], DSR [40], MDS [6] and MMCF [13]. However, the proposed DBN-CROFF method achieves a better and higher facer recognition rate from the LR image in the presence of an EYB dataset.

5.1.4 Performance of FERET dataset

Figure 7 illustrates the proposed performance of the DBN-CROFF face recognition based on the FERET dataset. More than 14,801 images from 1020 subjects are in the FERET dataset, and each subject has pictures with changes in illumination, expressions, and poses. Three probe sets and one standard gallery set are present in the dataset. The group of gallery set that comprises of 915 frontal face images is selected. The gallery set with one image per subject is used that contains 915 images by means of different expressions. Through selecting the training set, the efficiency of our proposed system is evaluated. The downsampling operation
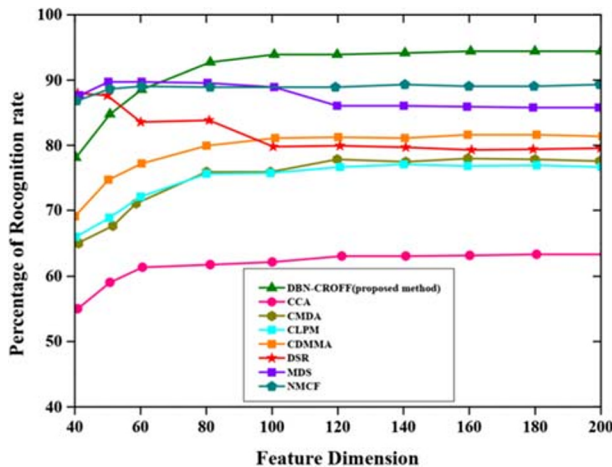

Fig. 6 DBN-CROFF facial recognition performance based on the EYB dataset
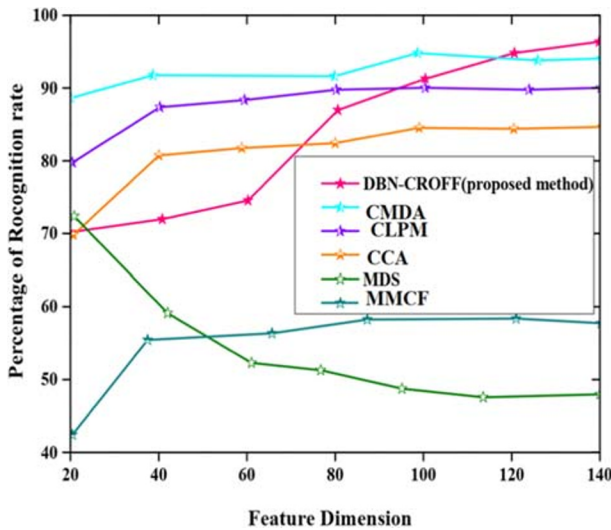
Fig. 7 Proposed DBN-CROFF face recognition performance based on the FERET dataset

generates a low-resolution probe image size 12 × 12. The proposed DBN-CROFF method is compared with existing CMDA [40], CCA [35], CLPM [18], MDS [6] and MMCF [13]. The proposed DBN-CROFF method despite these challenges surpasses the accuracy of face recognition.

5.2 Performance of face recognition rate with different sizes

The dimensionality reduction efficiency of the proposed LR-FR work is outlined in Fig. 8. In this experiment, we have selected facial width, cheeks size, skin tone, nose shape, and lip shape
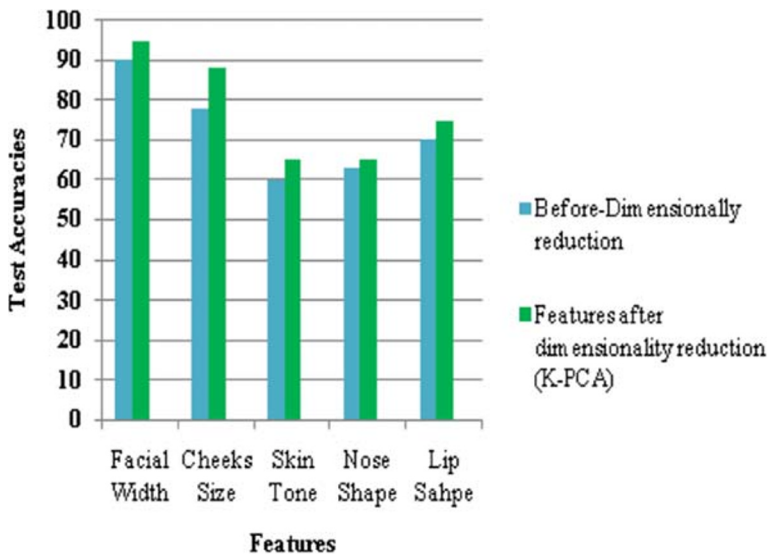

Fig. 8 Feature dimensionality reduction performance

feature. Particularly, the experiment is conducted between before feature dimensionality reduction and after dimensionality reduction. From Fig. 8, the proposed k-PCA method provides better feature dimensionality reduction results than before dimensionality reduction. Finally, the test accuracies of each feature after dimensionality reduction performance results are higher.

5.3 Performance of face recognition rate with different sizes

The face recognition performances of proposed DBN-CROFF with different image sizes are designed in Fig. 9. In this experiment, we have chosen different image sizes of LR images such as 8×8, 16×16 and 32×32 for face recognition from the LR image. The experiment is held between different image sizes with four datasets namely LFW, Multi-PIE, Extended Yale-B, and FERET. From this experiment, we examine that the proposed DBN-CROFF method with 32×32 image size accomplishes a higher face recognition rate than the other two image sizes.

5.4 Different probe resolution performance of proposed DBN-CROFF

The proposed face recognition rate based on the datasets with various probe image resolutions are depicted in Fig. 10. This experiment uses sample images of different resolutions for the proposed method, with and without a morphological pre-processing operation. Here, we used three different probe image resolutions such as 8×8, 12×12 and 16×16 with four datasets namely LFW, Multi-PIE, Extended Yale-B, and FERET. The respective optimal feature dimensions are selected for validation of the performance of the experiment at various probes resolutions with a fixed number of training samples per subject. Each of these dataset images shown better face recognition with optimal probe image resolution.

5.5 State-of-art results

The performance of the k-PCA method suggested for dimensionality reduction is described in Table 1. Initially, the facial width, cheeks size, skin tone, nose shape, and lip shape feature to convert the high dimensional features into low dimensional one. In this experiment, the various
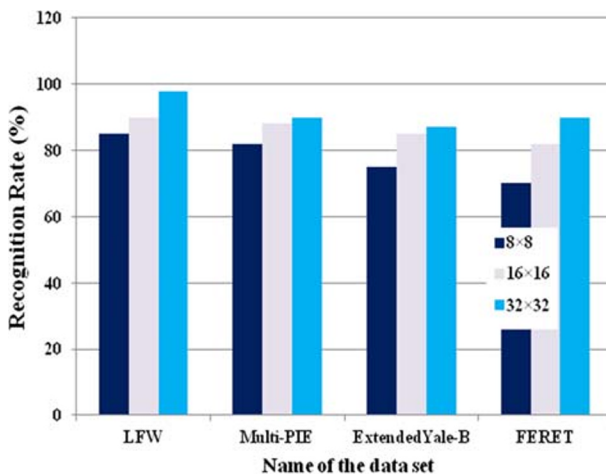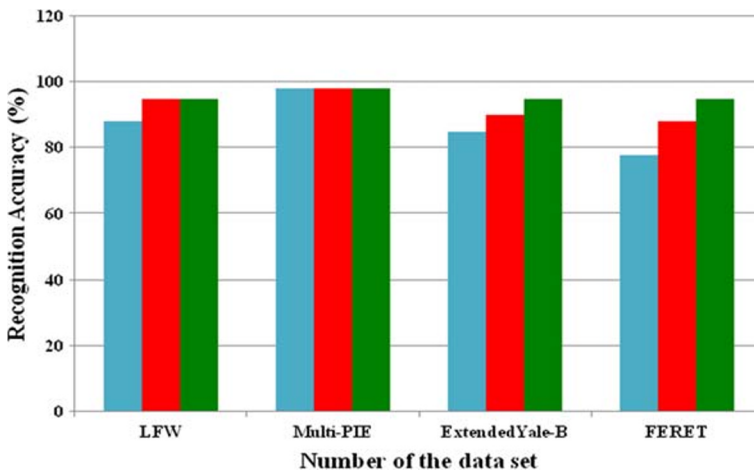

Fig. 9 Proposed face recognition rate with different image sizes

Fig. 10 Proposed face recognition rate based on the datasets with various probe image resolutions

dimension rates such as 5 and 7 are utilized to evaluate the performance of the conventional techniques. The DBN method is employed with k-PCA, t-Distributed Stochastic Neighbor Embedding (t-SNE), Locally Linear Embedding (LLE), and PCA. According to Table 1, the proposed k-PCA (DBN-CROFF) delivers better and accurate performance results than other existing methods in each feature.

Table 2 shows the state-of-the-art results for facial recognition rate accuracy. In this experiment, the proposed methodology(Proposed (DBN-CROFF)) is compared with three conventional techniques namely the Low-rank matrix, and sparse error matrix(LRM-SEM) [26], Multi-scale [8] and SVM [19] are used. The proposed (DBN-CROFF) extracted facial

Table 1 State-of-art performance with different feature dimensions

Method	Feature	Identification rate %	
		5 dimension	7 dimension
Kernel PCA + (DBN-CROFF)	HOG of facial width	97.85	97.98
	HOG of cheeks size	95.28	94.78
	HOG of skin tone	89.95	88.65
	HOG of nose shape	93.73	90.77
	HOG of lip shape	93.78	92.01
t-SNE + (GSA-DBN)	HOG of facial width	83.95	80.84
	HOG of cheeks size	70.68	72.52
	HOG of skin tone	62.54	53.63
	HOG of nose shape	68.86	72.53
	HOG of lip shape	70.12	60.02
PCA + (GSA-DBN)	HOG of facial width	59.99	56.48
	HOG of cheeks size	62.25	64.87
	HOG of skin tone	62.68	74.54
	HOG of nose shape	54.79	58.82
	HOG of lip shape	68.23	67.03
LLE + (GSA-DBN)	HOG of facial width	71.52	73.52
	HOG of cheeks size	58.78	60.57
	HOG of skin tone	72.58	78.92
	HOG of nose shape	67.56	50.84
	HOG of lip shape	72.67	66.53

Table 2 Performance comparison with conventional face recognition techniques in terms of accuracy

Methods	Extracted features	Name of the datasets	Accuracy (%)
Proposed (DBN-CROFF)	Facial width, cheeks size, skin tone, nose shape, and lip shape	FERET	93%
		EYB	95%
		M-PIE	90%
		LFW	89%
Low-rank matrix and sparse error matrix [26]	Facial width, mouth, and eyes	FERET (Fa)	89%
		FERET (BaBe)	91%
		Multi-PIE	97%
		LFW	88%
		Extended Yale-B	94%
Multi-scale [8]	Lip, eyebrow, cheek, eye and furrow	LFW	72%
		Multi-PIE	92%
		NUST-RWFR	67.7%
SVM [19]	Learned and Handicrafted	CK+	92.5%

width, cheeks size, skin tone, nose shape, and lip shape features. The FERET, Extended Yale-B(EYB), Multi-PIE(M-PIE), and LFW dataset provides 93%, 95%, 90%, and 89% accuracy rates. Also, facial width, mouth, and eye characteristics are extracted using LRM-SEM [26]. The FERET (Fa), FERET (BaBe), M-PIE, LFW, and EYB dataset provided 89%, 91%, 97%, 88%, and 94% accuracy rates. The existing Multi-scale [8] method with LFW, Multi-PIE, and NUST-RWFR datasets accomplished 72%, 92%, and 67.7% recognition accuracy rates. The existing SVM [19] method with CK+ dataset provided 92.5% accuracy. However, the proposed DBN-CROFF method provided higher recognition accuracy than other existing methods.

6 Conclusion

This paper proposed Deep Belief Network-Crossover based Firefly (DBN-CROFF) method for face recognition from LR images. Experimentally, the HR images were mainly down-sampled to obtain the images within a specific size similar to the LR images. LR face images are collected from four datasets namely LFW, Multi-PIE, Extended Yale-B, and FERET datasets. The proposed DBN-CROFF method with four datasets performances is evaluated using existing CMDA, CLPM, CDMMA, CCA, DSR, MDS, and MMCF methods. The k-PCA method yields higher feature dimensionality reduction results when compared to other t-SNE, LLE, and PCA methods. Finally, the proposed DBN-CROFF method with FERET, EYB, M-PIE, and LFW dataset provided 93%, 95%, 90%, and 89% accuracy rates.

References

1. Abebe HB, Hwang C-L (2019) RGB-D face recognition using LBP with suitable feature dimension of depth image. IET Cyber-Physical Systems: Theory and Applications 4(3):189–197

2. Abiantun R, Juefei-Xu F, Prabhu U, Savvides M (2019) SSR2: Sparse signal recovery for single-image super-resolution on faces with extreme low resolutions. Pattern Recogn 90:308–324
3. Abualigah LMQ (2019) Feature selection and enhanced krill herd algorithm for text document clustering (pp. 1–165). Springer, Berlin
4. Abualigah LM, Khader AT, Hanandeh ES (2018) Hybrid clustering analysis using improved krill herd algorithm. Appl Intell 48(11):4047–4071
5. Alenazy WM, Alqahtani AS (2020) Gravitational search algorithm based optimized deep learning model with diverse set of features for facial expression recognition. J Ambient Intell Humaniz Comput p 1–16
6. Biswas S, Bowyer K, Flynn PJ (2011) Multidimensional scaling for matching low-resolution face images. IEEE Trans Pattern Anal Mach Intell 34(10):2019–2030
7. Dornaika F, Davoine F (2008) Simultaneous facial action tracking and expression recognition in the presence of head motion. Int J Comput Vis 76(3):257–281
8. Gao G, Yu Y, Yang M, Huang P, Ge Q, Dong Y (2020) Multi-scale patch based representation feature learning for low-resolution face recognition. Appl Soft Comput 90:106183
9. Gross M, Cohn J, Kanade T, Baker S (2010) Multi-pie. Image Vis Comput 28(5):807–813
10. Haghighat M and Abdel-Mottaleb M (2018). "Low resolution face recognition in surveillance systems using discriminant correlation analysis," IEEE International Conference on Automatic Face & Gesture Recognition, pp. 912–917
11. Heinsohn V, Mery P (2019) Face recognition in low-quality images using adaptive sparse representations. Image Vis Comput 85:46–58
12. Hinton G, Osindero S, The Y-W (2006) A fast learning algorithm for deep belief nets. Neural Comput 8(7): 1527–1554
13. Huang H, He H (2010) Super-resolution method for face recognition using nonlinear mappings on coherent features. IEEE Trans Neural Netw 22(1):121–130
14. Huang, Gary, Marwan Mattar, Tamara Berg, and Eric Learned-Miller (2008). "Labeled faces in the wild: a database forstudying face recognition in unconstrained environments". Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition, Erik Learned-Miller and Andras Ferencz and Frédéric Jurie, Marseille
15. Jiang J, Hu R, Wang Z, Cai Z (2016) CDMMA: Coupled discriminant multi-manifold analysis for matching low-resolution face images. Signal Processing 124:162–172
16. Jolliffe I. T. (2011). "Principal Component Analysis," Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1094–1096
17. Kurup AR, Ajith M, Ramón MM (2019) Semi-supervised facial expression recognition using reduced spatial features and Deep Belief Networks. Neurocomputing 367:188–197
18. Li B (2009) Low-resolution face recognition via coupled locality preserving mappings. IEEE Signal processing letters 17(1):20–23
19. Li L, Lingqiao Qin XQ, Zhang J, Wang Y, Ran B (2019) Day-ahead traffic flow forecasting based on a deep belief network optimized by the multi-objective particle swarm algorithm. Knowledge-Based Systems 172: 1–14
20. Li L, Qin L, Qu X, Zhang J, Wang Y, Ran B (2019) Day-ahead traffic flow forecasting based on a deep belief network optimized by the multi-objective particle swarm algorithm. Knowl-Based Syst 172:1–14
21. Lu Z, Jiang X, Kot A (2018) Deep coupled resnet for low-resolution face recognition. IEEE Signal Processing Letters 25(4):526–530
22. Marciniak T, Chmielewska A, Weychan R, Parzych M, Dabrowski A (2015) Influence of low resolution of images on reliability of face detection and recognition. Multimed Tools Appl 74(12):4329–4349
23. Pan X, Li X (2019) A new and efficient firefly algorithm for numerical optimization problems. Neural Comput & Applic 31(5):1445–1453
24. Phillips J, Moon H, Rizvi S, Rauss PJ (2020) The FERET evaluation methodology for face-recognition algorithms. IEEE Trans Pattern Anal Mach Intell 22(10):1090–1104
25. Rejeesh MR (2019) Interest point based face recognition using adaptive neuro fuzzy inference system. Multimed Tools Appl 78(16):22691–22710
26. Saad S, Lam K-M, Lai S-C (2019) Learning sparse discriminant low-rank features for low-resolution face recognition. J Vis Commun Image Represent 63:102590
27. Sun J, Shen Y, Yang W, Liao Q (2020) Classifier shared deep network with multi-hierarchy loss for low resolution face recognition. Signal Process Image Commun pvol82:115766
28. Sundararaj V (2016) An efficient threshold prediction scheme for wavelet based ECG signal noise reduction using variable step size firefly algorithm. Int J Intell Eng Syst 9(3):117–126
29. Sundararaj V (2019) Optimal task assignment in mobile cloud computing by queue based ant-bee algorithm. Wirel Pers Commun 104(1):173–197

30. Sundararaj V, Anoop V, Dixit P, Arjaria A, Chourasia U, Bhambri P, Rejeesh MR, Sundararaj R (2020) CCGPA-MPPT: Cauchy preferential crossover-based global pollination algorithm for MPPT in photovoltaic system. Prog Photovolt 28 (11):1128–1145
31. Ullah Haq, Khattak Khan and Mahmood Z (2019). "A Robust Face Recognition Method for Occluded and Low-Resolution Images," International Conference on Applied and Engineering Mathematics (ICAEM), pp. 86–91
32. Vinu S (2019) Optimised denoising scheme via opposition-based self-adaptive learning PSO algorithm for wavelet-based ECG signal noise reduction. Int J Biomed Eng Technol 31(4):325
33. Vinu S, Muthukumar S, Kumar RS (2018) An optimal cluster formation based energy efficient dynamic scheduling hybrid MAC protocol for heavy traffic load in wireless sensor networks. Comput Secur 77:277–288
34. Wahid F, Ghazali R, Ismail LH (2019) Improved firefly algorithm based on genetic algorithm operators for energy efficiency in smart buildings. Arab J Sci Eng 44(4):4027–4047
35. Wang Z, Yang W, Ben X (2015) Low-resolution degradation face recognition over long distance based on CCA. Neural Comput & Applic 26(7):1645–1652
36. Xu Haiyan, and Konstantinos Plataniotis (2016). "Affective states classification using EEG and semi-supervised deep learning approaches," International Workshop on Multimedia Signal Processing (MMSP), pp.1–6
37. Yi D, Lei Z, Liao S and Li SZ (2014). "Learning face representation from scratch," arXiv preprint arXiv, pp.1411.7923
38. Yin B, Guo Z, Liang Z, Yue X (2018) Improved gravitational search algorithm with crossover. Comput Electr Eng 66:505–516
39. Zangeneh E, Rahmati M, Mohsenzadeh Y (2020) Low resolution face recognition using a two-branch deep convolutional neural network architecture. Expert Systems with Applications 139:112854
40. Zou W, Yuen P (2011) Very low resolution face recognition problem. IEEE Trans Image Process 21(1): 327–340