Optimization Technique to Linear Discriminant Regression for Face Recognition

T.Syed Akheel[1], Dr.V.Usha Shree[2] ,Dr.S.Aruna Mastani[3]

[1*] Research Scholar , Dept. of Electronics and communication Engg., JNTUA,Ananthapuramu,

[2] Supervisor & Professor, Dept. of Electronics and communication Engg, Joginapally B.R Institute of Engineering &Technology, Hyderabad,

[3]Co-Supervisor & Asst.Professor, Dept. of Electronics and communication Engg, JNTUCE, Ananthapuramu,

***Abstract*—** Facial recognition from picture and video stream is gaining popularity in the field of biometrics due to its variety of applications. The performance of facial recognition systems can be considerably improved by modelling the key parameters of the reconstruction error. In this paper we propose a linear collaborative discriminant regression classification system with cyclic exploration whale optimization (CEWO) and analyses its performance metrics. The proposed system makes use of collaborative between-class reconstruction error and within-class reconstruction error. This system optimizes the parameters using cyclic exploration whale optimization such that it minimizes the within-class reconstruction error and at the same time maximizes between-class reconstruction error. Earlier research using regression classification considered minimizing the within class errors or maximizing the between class error. Our system which utilizes both the within class and between class errors improves the recognition rate when compared with different conventional optimization techniques.

Keywords—Face Recognition; Active Appearance Model; Linear collaborative Discriminant regression classification; Whale Optimization;

I. INTRODUCTION

Facial recognition is to detect a person in the image or video, from the known dataset on which the system is developed. It depends on complex learning mechanisms, which depend on complex calculations [1]. With the evolution of machine learning techniques, the facial recognition has taken a big leap in the advancement [2]. But this advancement poses multiple challenges like training the system on processed databases, which extract the main features of the facial images. These features have to be aligned with the testing image features, which is a challenging task as the testing images could have different orientations and various light conditions and various items covering the face like the sunglasses, hat and headbands etc [3].

Different optimizing algorithms are proposed by researchers in the past, taking one or more parameters like reconstruction errors into consideration, yet yielding low recognition rates when tested on different subjects [4]. So, a facial recognition system developed on key features of the database is required to cater various applications. Here a novel facial recognition system is developed on regression classification, optimized using cyclic exploration whale optimization technique. The performance analytics are compared with the existing regression and optimization techniques.

Principle component analysis (PCA) is one of the most well-known methods for feature extraction. PCA aims to transform the original data into a low-dimensional subspace where the variance of the data is preserved as much as possible. PCA is an effective data representation technique, but it may be unsuitable for classification since PCA cannot discover the discriminant structure of the data.

In view to improve the better accuracy regression methods has been effectively used to improve the results, regression method with modular approach is presented [7] but still the LRC and other regression methods do not take between classes into consideration. Numerous approaches are there for recognizing and detecting the face. Feature extraction approach can be done in which, the extracted features from the face can be processed and compared , for obtaining the better solution the optimization techniques are presented for better and accurate results , here the Enhanced WOA [10] is compared to state of art optimization algorithms like GA [15] ,ABC[14], PSO[13] ,GWO[12] and FF[11].

II. LITERATURE REVIEW

A. Related Works

Xiaochao Qu et al [1] have presented linear collaborative discriminant regression classification for face recognition. The idea of collaboration representation of many classes is usually compared with reconstruction error. This paper adopts a better between-class reconstruction error measurement which is obtained using the collaborative representation instead of class-specific representation. The main disadvantage of LCDRC is that it is used the single linear regression model which is consist of one predictor that leads to anomalous results in accuracy. S. M. Huang and J. F. Yang [2] have presented linear regression classification methodology with the help of class-specific representation where it was distinguished by Between-Class Reconstruction Error (BCRE) and Within-Class Reconstruction Error (WCRE) to find a discriminant subspace by maximizing the value of BCRE and minimizing the value of WCRE simultaneously. The main disadvantage of the LDRC is maximization of the overall between-class reconstruction error is easily dominated by some large class-specific between-class reconstruction errors, which makes the following LRC erroneous.

Xiaochao Qu et al [3] have presented an enhanced discriminant linear regression classification (EDLRC) algorithm to further improve the discriminant power of LDRC. They haven't used all those classes for calculating BCRE rather than they have only considered about the classes with small reconstruction error. Through maximizing the construction error of the true class's similar classes, their EDLRC increased the discriminatory power of LDRC. Their experiment showed that EDLRC performed better than LRC and LDRC for ORL and AR database.

I.Nassem et al [4] proposed a Linear Regression Classification (LRC) algorithm falls in the category of nearest subspace classification. The algorithm is extensively evaluated on several standard databases under a number of exemplary evaluation protocols reported in the face recognition literature. A comparative study with state-of-the-art algorithms clearly reflects the efficacy of the proposed approach. For the problem of contiguous occlusion, propose a Modular LRC approach, introducing a novel Distance-based Evidence Fusion (DEF) algorithm.

Zeng et al [5] focussed on the automatic affect recognition using visual and auditory modalities. The discussion will present a model which helps is presenting the best automatic model for face recognition.

GunSet al [6] highlight the continuity aspect in terms of input and output of images and provides the dimensionality of representation to higher level. Here the analysis and the response provided by the automatic system are also envisioned to be continuous over the course of time, within the boundaries of digital machine output.

Yong Xu et al.[7] have developed a sparse representation approach for face recognition using l2 regularization in which a noticeable performance can be attained in face recognition. The discriminative presentation could be attained by correlation reduction of presentations. In sparse representation approaches, the presentations correlation of the test sample generated from various classes could be minimized in terms of sparsity and collaboration.

Yuwu Lu *et al.* [8] have developed a new linear regression kernel for face recognition. The limitation of the linear regression is exploited, and the intersections of class specific samples are classified properly, and a new kernel linear regression model is developed which maps the data more linearly and performs better recognition in the low light conditions. This is then compared with three other models and proved that this system outperforms the other systems.

III. Design representation: Architecture and its subsequent phases

Appearance Model: The construction of full model of face image needs both the texture and shape model. Thus, the construction of numerical texture model is the next step, where the texture samples' alignment is needed to a reference texture frame and the appearance is composed of texture information. As the structuring of statistical appearance model is need to warp the colour channels, initially the control points are matched to the mean shape. The piecewise affine warping (i.e., Dividing the convex outlines of the mean shape by triangle sets) is done for matching the texture. The appearance model, $AP(\bar{x})$ is attained by concerning PCA to texture vectors as defined in Eq. (1), where AP_0 denotes the mean appearance vector, δ_i indicates the appearance parameter, and the synthesised appearance vector from affine wrapping is indicated by $AP_i(\bar{x})$.

$$AP(\bar{x}) = AP_0(\bar{x}) + \sum_{i=1}^{mm} \delta_i AP_i(\bar{x}) \quad (1)$$

The output of AAM based feature extraction is indicated as F that includes both shape and texture features.

A. Linear Regression Classification (LRC) [2]

Let denote the training face images of the ith class as $X_i \in \partial^{m \times n_i}$. Each column of X_i are n_i training face images, and $1 \le i \ge c$ where c is the total number of classes.

Assume y be the probe face image that can be represented using X_i according to

$$y = X_i \alpha_i \text{, where } 1 \le i \ge c \quad (2)$$

$\alpha_i \in \partial^{n_i \times 1}$ is the regression parameters; α_i can be calculated using the least-square estimation as,

$$\hat{\alpha}_i = \left(X_i^T X_i\right)^{-1} X_i^T y \text{, } 1 \le i \ge c \quad (3)$$

The reconstruction of y by each class can be obtained as,

$$\hat{y}_i = X_i \hat{\alpha}_i = X_i \left(X_i^T X_i\right)^{-1} X_i^T y = H_i y \text{, } 1 \le i \ge c \quad (4)$$

Where H_i is called hat matrix that maps y into $\hat{y}$ the reconstruction error of each class is calculated as

$$e_i = \|y - \hat{y}_i\|_2^2, \ \ldots .\ 1 \le i \ge c \qquad (5)$$

B. LDRC-based Classification[2]

Let Z belongs to $K^{n \times d}$ denote the projection matrix then within class and between classes can be obtain

$$E_W = \frac{1}{n}\sum_{i=1}^{n}(x_i - x_i^{intra})(x_i - x_i^{intra})^T$$
$$E_B = \frac{1}{n(c-1)}\sum_{i=1}^{n}\sum_{j=1, j \ne l_i}^{c}(x_i - x_{ij}^{inter})(x_i - x_{ij}^{inter})^T \qquad (6)$$

$$J(P) = \underset{P}{MAX}\ \frac{tr(Z^T E_b Z)}{tr[Z^T(E_W + \in I)Z} \qquad (7)$$

Where l_i the label of the sample x_i, x_i^{intra} the intra class reconstruction error and x_{ij}^{inter} is the inter class reconstruction vector of x_i with respect to j^{th} class, where $\in$ >0, $\in I$ is added to avoid the singularity of the matrix. The optimal projection of the matrix LDRC is composed by Eigen vectors and associated with first d largest values, This can be solve by identifying the largest d Eigen values

$$E_b.a_k = \lambda_k(E_W + \in I)a_k,\ k = 1,2....d \qquad (8)$$

C. LCDRC-based Classification[1]

Let the training matrix is represented as $X = \left[X_1,.....X_i,.....X_p\right] \in \Re^{p \times q_i}$, where p indicates the dimension of each training face image, q_i refers to the count of training face image from class i. Let the subspace projection matrix that is to be learned is denoted as $P \in \Re^{p \times d}$ and $d < p$. Each x_{ij} can be mapped to learned subspace by $g_{ij} = P^T x_{ij}$, where $1 \le j \le q_i$. The whole training face image matrix is then mapped as $G = P^T X \in \Re^{d \times q}$ and for each class $G_i = P^T X_i \in \Re^{d \times q_i}$. The collaborative between-class reconstruction error (CBCRE) and within-class reconstruction error (WCRE) are defined as in Eq. (9)

$$CBCRE = \frac{1}{q}\sum_{i=1}^{c}\sum_{j=1}^{q_i}\| g_{ij} - \hat{g}_{ij}^{inter}\|_2^2$$
$$WCRE = \frac{1}{q}\sum_{i=1}^{c}\sum_{j=1}^{q_i}\| g_{ij} - \hat{g}_{ij}^{intra}\|_2^2 \qquad (9)$$

where $\hat{g}_{ij}^{inter} = G_{ij}^{inter}\alpha_{ij}^{inter}$ and $\hat{g}_{ij}^{intra} = G_{ij}^{intra}\alpha_{ij}^{intra} \cdot G_{ij}^{inter}$ is the G with G_i eliminated and G_{ij}^{intra} is the G_i with g_{ij} eliminated. α_{ij}^{inter} and α_{ij}^{intra} is attained by Eq. (10).

$$\hat{\alpha}_i = \left(X_i^T X_i\right)^{-1} X_i^T g, i = 1,2,...c \qquad (10).$$

Prior obtaining the P , the value of α in learned subspace is unknown for us. However, the $\hat{\alpha}$ is evaluated in the original space and $\hat{\alpha}$ is used as the approximation of α. From Eq. (9), the difference among CBCRE and BCRE is seen, and CBCRE uses cross-class collaborative representation and BCRE uses class-specific representation. As per the relationship among X and G, CBCRE and WCRE is rewritten as in Eq. (11). This is again rewritten as in Eq. (12).

$$CBCRE = \sum_{i=1}^{c}\sum_{j=1}^{q_i}\| P^T x_{ij} - P^T X_{ij}^{inter}\alpha_{ij}^{intra}\|_2^2$$
$$WCRE = \sum_{i=1}^{c}\sum_{j=1}^{q_i}\| P^T x_{ij} - P^T X_{ij}^{intra}\alpha_{ij}^{intra}\|_2^2 \qquad (11)$$

The Eigen vectors EI_b and EI_w is determined as in Eq. (12).

$$EI_b = \frac{1}{q}\sum_{i=1}^{c}\sum_{j=1}^{q}\left(x_{ij} - X_{ij}^{inter}\alpha_{ij}^{inter}\right)\left(x_{ij} - X_{ij}^{inter}\alpha_{ij}^{inter}\right)^T$$
$$EI_w = \frac{1}{q}\sum_{i=1}^{c}\sum_{j=1}^{q_i}\left(x_{ij} - X_{ij}^{intra}\alpha_{ij}^{intra}\right)\left(x_{ij} - X_{ij}^{intra}\alpha_{ij}^{intra}\right)^T \qquad (12)$$

This can be solve by identifying the largest d eigen values and the according eigen values as the following Eq. (17), where $\lambda_1 \ge ..... \ge \lambda_k .....\lambda_d$ and $P = \left[p_1,....p_k,...,p_d\right]$. MMC solves the small sample size problem (SSP), where the face image dimensions is greater than the count of training face images.

$$\left(EI_b - EI_w\right)P_k = \lambda_k p_k, k = 1, 2....d \qquad (13)$$

IV. PROPOSED ALGORITHM FOR PROJECTION MATRIX OPTIMIZATION

A. Fitness Function

To evaluate the function at first the error is calculated i.e $error$ between actual value, act and predicted value, $pred$ is evaluated followed by fitness function. In Eq. (13), λ indicates the regularization and the minimization of overall error along with λ is consider as the major intensive of this proposed work .The analysis is presented in Fig 1 by varying the regularization constant the parameters are calculated for showing the improvement in the optimization technique.

$$error = \left(act - pred\right) \qquad (14)$$

$$FT = Min\left(Sum\left(error\right) + \lambda * \sum_{i=1}^{NU}\left(P\right)^2\right) \qquad (15)$$

B. Whale Optimization Algorithm[10]

Whale optimization algorithm was proposed by Seyedali Mirjalili and Andrew Lewis in 2016, Different phases are there in WOA algorithm [10], and they are Encircling prey, Spiral bubble-net feeding maneuver, and search for prey. An improvement is made with the conventional WOA to get the optimal P , which is explained below.

Encircling prey: Eq. (16) defines the position update in the direction of the best search agent, where tn indicates the current iteration, $\vec{K}$ and $\vec{L}$ represents the coefficient vectors, p^* indicates the position vector of best solution, position vector is indicated by $\vec{P}$, $|\,|$ is the absolute value, and it denotes the element by element multiplication. Eq. (18) and (19) shows the evaluation of $\vec{K}$ and $\vec{L}$

$$\vec{H} = |\vec{L}.\vec{P}^*(tn) - \vec{P}(tn) \qquad (16).$$

$$\vec{P}(tn+1) = \vec{P}^*(tn) - \vec{K}\cdot\vec{H} \qquad (17)$$

$$\vec{K} = 2.\vec{a}..\overrightarrow{ru} - \vec{a} \qquad (18)$$

$$\vec{L} = 2.\overrightarrow{ru} \qquad (19)$$

In the exploration and exploitation phase, $\vec{a}$ is linearly reduced from 2 to 0, $\overrightarrow{ru}$ denotes the random vector in [0,1].

C.Enhanced Whale Optimization Algorithm[EWOA]

This paper proposes a Enhanced WOA algorithm for attaining the optimal projection matrix, P , which is the improvement of conventional WOA algorithm The enhancement is made in case of no improvement in Fitness evaluation. This paper fixes N_{cycle}=3, and finds the best solution as the optimal projection matrix P^*.

V.RESULTS AND DISCUSSIONS

A.Simulation Setup

The proposed method i.e. Cyclic exploration whale optimization (CEWO) simply Enhanced Whale optimization algorithm (EWOA) optimization technique provide the better and best solution in comparison with other techniques, Simulation results are performed using the image database downloaded from the link shown which is freely available for from URL: http://cswww.essex.ac.uk/mv/allfaces/index.html database and is implemented on MATLAB 2018b to validate the efficiency and other parameters.. The database includes both male and female images with different variations. The proposed method was compared to other conventional methods like conventional Whale Optimization Algorithm (WOA) [10], Grey Wolf Optimization (GWO) [11], FireFly (FF) [12], Particle Swarm Optimization (PSO) [13], Artificial Bee Colony (ABC) [14], and Genetic Algorithm (GA) [15]. The performance of proposed model was analyzed in terms of measures like Fig 1(a) shows the Recognition rate which shows the recognition rate performance over other methods, Fig 1 (b) graph for Sensitivity shows its improvement, Fig 1(c) graph which shows that the performance of specificity with respect to percentage of learning has improved compared to other conventional methods, Fig 1(d) graph for the improvement in Precision.

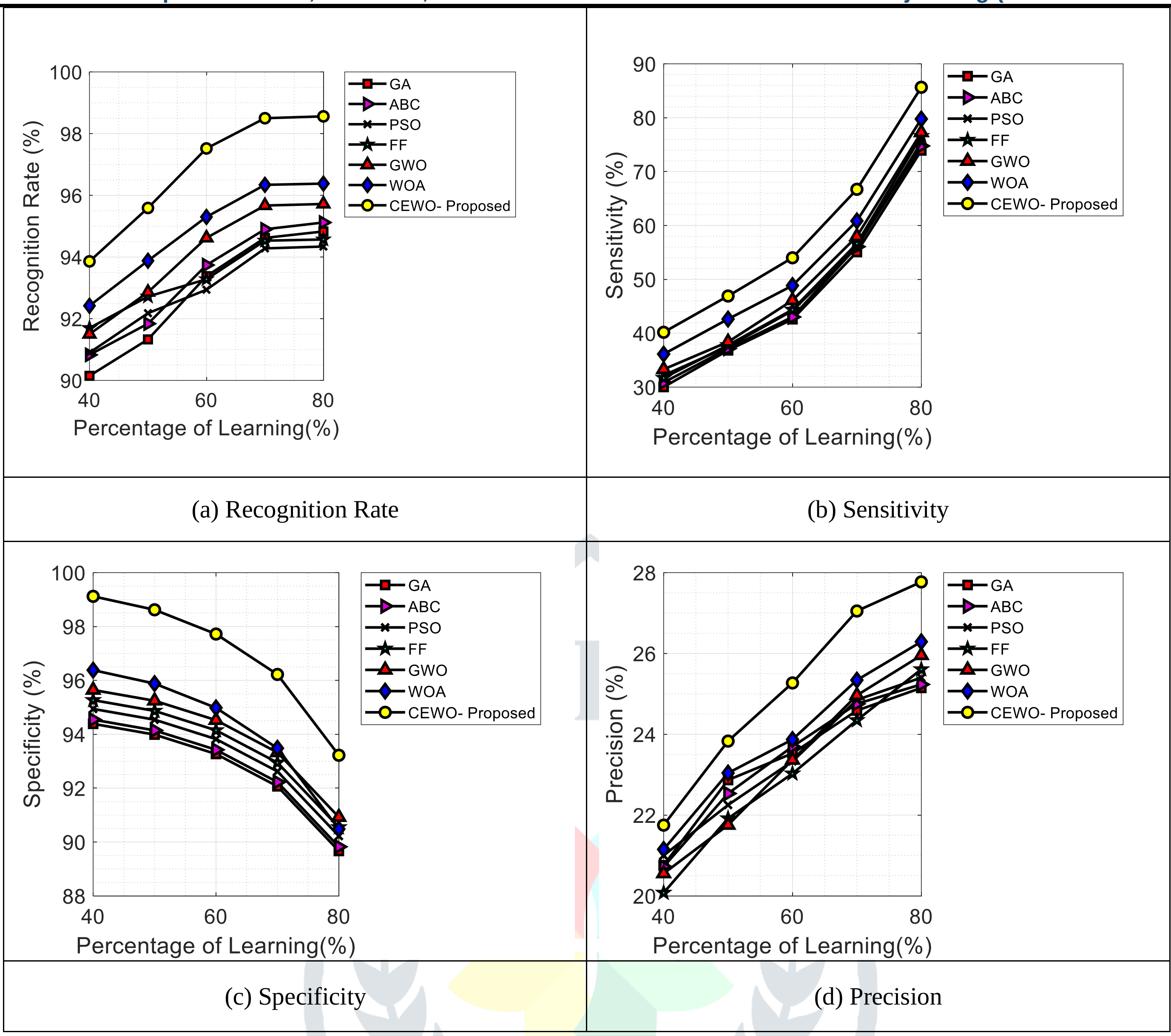

Fig 1. Parameters on proposed and conventional face recognition model (a) Recognition rate (b) Sensitivity (c) Specificity (d) precision

Table 1. Performance analysis metrics of the proposed model

Measures (%)	FACE 94 Dataset	ORL Dataset
Recognition rate	0.989	0.9688
Sensitivity	0.8157	0.6222
Specificity	0.9352	0.9903
Precision	0.2685	0.6222

VI.CONCLUSION

In this paper a new regression classification system, with cyclic exploration whale optimization technique is proposed, which considers the reconstruction errors of both the within class and between class datasets. By combining both reconstruction errors to optimize the system gives better performance than the conventional algorithms. The performance metrics shows that the current system has better recognition rate than the existing algorithms. Further the system could be integrated with advanced image processing techniques for the complex datasets, where it contains multiple dimensional irregularities.

VII.REFERENCES

[1] Xiaochao Qu, Suah Kim, Run Cui, Hyoung Joong Kim, “Linear collaborative discriminant regression classification for face recognition”, Journal of Visual Communication and Image Representation, Vol. 31, pp. 312–319, 2015.
[2] Shih-Ming Huang and Jar-Ferr Yang "Linear Discriminant Regression Classification for Face Recognition", IEEE Transactions On Signal Processing Letters, Vol. 20, No. 1, 2013
[3] Xiaochao Qu, Hyoung Joong Kim, "Enhanced Discriminant Linear Regression Classification for Face Recognition", IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP) Symposium on Cognitive Computing in Information Processing, PP.1-5, Singapore, 2014

[4] I. Naseem, R. Togneri, and M. Bennamoun, “Linear regression for face recognition,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 32, no. 11, pp. 2106–2112, 2010.

[5] Z.Zhang,M.Pantic,G,I Roisaon et Al A Survey On Face Recogntion Models,IEEE Transancations Vol31, Pp39-58,Jan 2009

[6] H. gunes and schuller ."Categoricl And Dimesnion Affect Analysis In Continous Input ,imagevis compute vol 31 . Volume 31, Issue 2, February 2013, Pages 120-136

[7] Y. Xu, Z. Zhong, J. Yang, J. You and D. Zhang, "A New Discriminative Sparse Representation Method for Robust Face Recognition via l2 Regularization," IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 10, pp. 2233-2242, Oct. 2017.

[8] Lu, Yuwu, Xiaozhao Fang, and Binglei Xie. "Kernel linear regression for face recognition." Neural Computing and Applications 24, no. 7 (2014): 1843-1849.

[9] Zhou, Changjun, Lan Wang, Qiang Zhang, and Xiaopeng Wei. "Face recognition based on PCA and logistic regression analysis." Optik 125, no. 20 (2014): 5916-5919.

[10]Seyedali Mirjalili and Andrew Lewis, " The Whale Optimization Algorithm", Advances in Engineering Software, vol. 95, pp.51–67, 2016.

[11]Seyedali Mirjalili, Seyed Mohammad Mirjalili and Andrew Lewis, " Grey Wolf Optimizer", Advances in Engineering Software, vol.69, pp.46–61, 2014

[12]A.H. Gandomi , X.-S. Yang , S. Talatahari and A.H. Alavi, " Firefly algorithm with chaos", Commun Nonlinear Sci Numer Simulat, vol. 18, pp. 89–98, 2013.

[13]M.E.H.Pedersen and A.J.Chipperfield, " Simplifying Particle Swarm Optimization", Applied Soft Computing, vol. 10, no. 2, pp. 618-628, 2010.

[14]D.Karaboga and B.Basturk, " On the performance of artificial bee colony (ABC) algorithm", Applied Soft Computing, vol. 8, no. 1, pp. 687-697, 2008.

[15]JohnMcCall, "Genetic algorithms for modelling and optimization", Journal of Computational and Applied Mathematics, vol. 184, no. 1, pp. 205-222, 2005