2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE) | 978-1-6654-3897-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICECCE52056.2021.9514139

Face Recognition System using Histograms of Oriented Gradients and Convolutional Neural Network based on with Particle Swarm Optimization

Sulayman Ahmed
ENET'COM
Universite de Sfax
Sfax, Tunisia
sulaimanah2000@uokirkuk.edu.iq,

Mondher Frikha
ENET'COM
Universite de Sfax
Sfax, Tunisia
mondher.frikha@enetcom.usf.tn,

Taha Darwassh Hanawy Hussein
Kirkuk University
Kirkuk, Iraq
taha_hussein@uokirkuk.edu.iq,

Javad Rahebi
Department of Software Engineering
Istanbul Ayvansaray University
Istanbul, Turkey
cevatrahebi@ayvansaray.edu.tr

***Abstract*— In this paper, Histograms of Oriented Gradients dependent on the strong point of convolutional neural organization which is new methodology for evenness face data set, is introduced. A proposed face acknowledgment framework was created to be utilized for various purposes. We utilized Gabor wavelet change for include extraction of evenness face preparing information and afterward we utilized profound learning technique for acknowledgment. We executed and assessed the proposed strategy on ORL and YALE data sets with Matlab 2020b. Besides, similar trials were directed applying Particle Swarm Optimization (PSO) for include determination approach. The execution of Gabor wavelet include extraction with a high number of preparing picture tests has end up being more viable than different strategies in our examination. The acknowledgment rate while carrying out the PSO strategies on ORL data set is 86.62% while it is 92.6% with the three techniques on YALE data set. In any case, the utilization of PSO calculation has expanded the exactness rate to 95.88% for ORL information base and 95.23% on YALE data set.**

Keywords*— *Convolutional neural network, Face recognition, Histograms of Oriented Gradients

I. INTRODUCTION

Face recognition (FR) is a biometric technique that belongs to biometrics, which uses physiological characteristics such as (the face, fingerprint, eye, etc.), and anatomy characteristics such as (signature, walking, etc.). These characteristics are special for each person to identify the identity of individuals and available to all people, unique to each person, and are permanent and difficult to emulate or falsify, unlike other traditional techniques such as passwords and identity cards that can be stolen or used fraudulently by others [1].

For the human brain, face recognition is considered as a high-level visual work where a person can detect and identify faces in a scene in an easy way, on another hand, developing a computer system that performs these tasks is a major challenge that hides many complex processes. This challenge becomes even greater because of some factors that may hinder facial recognition, especially when the conditions for getting a facial image are inappropriate and uncontrollable. There are two types of facial differences: one resulting from the physical similarity between individuals and this type is generally limited. By contrast, the changes in the person itself are much greater and can be attributed to several factors: facial expressions, lighting conditions, etc., which may cause the face recognition system to fail. In general, several scientific researchers have tried to solve these problems, but they are not perfect yet [2][3].

The face, certainly, is a biological characteristic used by humans to identify and know each other. The purpose of designing a face recognition system is to simulate the human recognition system and has been extensively studied [4][5][6]. However, these studies are not effective in the case of a problem of small sample size, and thus the results are still modest, this study presents a solution to this problem.

II. RELATED WORKS

In literature, there are several methods used for face recognition such as holistic, local or hybrid methods [7][8]. Then again, late examination uncovered that a balance based dataset for face acknowledgment is a helpful strategy to take care of the face acknowledgment issue; consequently, it is feasible to acknowledge face acknowledgment utilizing the evenness property of the face [9]. The balance property of the face is exceptionally valuable to take care of two principle actually open issues in face acknowledgment, first, the set number of face preparing tests, and second, the varieties in postures and outward appearances and also enlightenment conditions. The proposed strategy utilizes the balance property of the face to lessen the impact of those two issues. The FR system, in general, consists of three main step; the preprocessing step, the feature extraction step and the classification step [10][11]. The preprocessing step stabilizes and prepares the input for feature extraction step so that good features can be extracted from the image [12]. Accordingly, an ideal preprocessing is supposed to eliminate the irrelevant information (e.g., illumination, background, rotation) [11]. The preprocessing step is used by many authors in their studies [11][10][13][14], since they use many different techniques such

978-1-6654-3897-1/21/$31.00 ©2021 IEEE

as Gaussian filter [15], Difference of Gaussian (DOG) [13], 2-D Wavelet Transform (2D WT) [16][17], Gabor filter [10], gamma correction [13] and Histogram Equalization (HE) [18]. The feature extraction's goal is to extract relevant features from a face image. However, extracting information from the image is easy for a human but it is a very challenging task for computer vision [19]. There exist many algorithms that are used for feature extraction including Local Binary Pattern (LBP) [19], Gray Level Co-occurrence Matrix GLCM [20], Gabor [11], Speeded Up Robust Features (SURF) [21], Scale Invariant Feature Transform (SIFT) [22] and Linear Discriminant Analysis (LDA) [23]. The classification step is the last and actual recognition process. It is used to match the feature vector of the test image that is obtained from the feature extraction step with the feature vectors of the train images which are already stored in a database. There are many techniques that are used for classification [24] and they vary from the Euclidean Distance [25], K-Nearest Neighbor [19], Support Vector Machine (SVM) [26] to advanced algorithms like Neural Networks [27][28]. Although there are some proposed methods to overcome these problems using the property of symmetry in the face, such problems are still considered open that are not solved yet. One of these recent methods is proposed by the authors of [9], since they improve the rate of accuracy by training their FR system using data obtained from original and symmetrical face images.

III. Proposed Method

A. Feature extraction using Histograms of Oriented Gradients

Very popular method for feature extrication is the Histograms of Oriented Gradients (HOG) [29]. It is one type of descriptors that is used a lot in the human detection [18]. The HOG concept is to compute the gradient orientation and the gradient direct magnitude. To obtain the HOG of an image, first, the changes in X and Y are computed, then the magnitude and direction are obtained.

1) Computing Gradients

The main operation of HOG is the derivative, or the center difference, since, there are two derivatives, the x derivative and the y derivative, once these derivatives are obtained, the gradient magnitude and the gradient orientation can be computed.

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x-h)}{2h} \quad (1)$$

The magnitude is given by:

$$s = \sqrt{s_x^2 + s_y^2} \quad (2)$$

And the orientation is given by:

$$\theta = \arctan\left(\frac{s_y}{s_x}\right) \quad (3)$$

2) Blocks and Cells

Figure 1 shows a face image, it is assumed that this image is a 64x128 image, if this image is divided to 128 cells, then some blocks are taken, for example the first block is block 1 with 2x2 cells, then the second block will be 50 % overlapped, which block 2, so, each block will consist of 2x2 cells with size 8x8 which means 16x16, with 7x15 = 105 blocks in total.

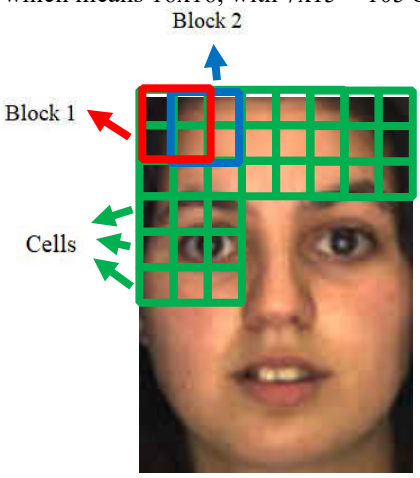


Figure 1. The blocks and the cells

3) HOG Feature Extraction Steps

To calculate the HOG for an image with 64 x 128, for example, the image is divided onto 16x16 blocks with 50% overlap, so therefore there will be 7x15 with total of 105 of blocks, and each block will consist of 2x2 cells, and the size will be 8x8, then the HOG is quantized with 9 directions or bins, if the direction is not one of the bins then some kind of interpolation can be done, also, the Gaussian can be applied to smooth the histogram, then all the descriptors can be concatenated since there are 105 of these block and each one is 9 dimensional, this will give a very large described, about 3780 dimension descriptor and this for the whole image of the block in the image.

B. Feature selection with Particle Swarm Optimization

Molecule Swarm Optimization (PSO) or known as the bird swarm calculation, was at first made in 1995 by Kenny and Eberhart [61]. PSO is a numerical technique that attempts to take care of the improvement issues. For every issue, there are particles (arrangement) fly over the trouble spot dependent on some numerical estimations for the speed and position of the molecule. Every molecule has wellness esteems that are estimated by the wellness capacity to be upgraded and has speed that manages the flying of the particles [62].

In computational procedures, PSO is utilized as an irregular advancement calculation for include choice and characterization. This is finished by iteratively attempting to choose the most family member and helpful arrangement of highlights to improve or keep up the grouping execution for a hearty facial acknowledgment framework [63].

The essential thought behind this calculation is; co-evolvement of various classes of birds instead of zeroing in on a specific class of birds, will contribute powerful hunt capacities [64].

First; every one of the particles are doled out with essential qualities, after that fit qualities for every particles are assessed. At that point the current fit worth is resolved, assuming it is superior to the past one, we update it to the current worth, however in the event that the old fit worth is better; we keep it [65]. The calculation closures and this interaction is rehashed until the best arrangement is gotten.

C. Convolutional neural network

The principle segment of a convolution neural organization (CNN) is the convolution layer. The methodology behind a convolution layer is that, an element which has been adapted locally for some random contribution (for instance any 2-D pictures) ought to be useful in different areas of that equivalent information source. For instance, an element for edge identification which was demonstrated valuable in one piece of the picture, may be useful in different districts of the picture at a potential general component extraction stage. The learning of different highlights in a picture like edges situated at a point or bends, is acquired by sliding the channels across the picture with a stage or step size which is consistent for a given convolution layer.

A CNN contain numerous convolutional and subsampling layers alternatively followed by completely appended layers. The contribution to a convolutional layer is a m × m × r picture where m is the tallness and width of the picture and r is the quantity of channels, for example a RGB picture has r=3. The convolutional layer will have k channels (or portions) of size n × n × q where n is more modest than the element of the picture and q can be something very similar or more modest as the quantity of diverts and change in each center.

IV. RESULTS AND DISCUSSION

This section of the paper illustrate the experimental results that are derived from the simulation using MATLAB software. The proposed method uses three steps for the face recognition. At first step the feature extraction method uses the HoG feature extraction and save in the variable. In second step that is the feature selection step, the Particle Swarm Optimization (PSO) is used to find the best features from the HoG method. Normally the feature numbers are high and with using the PSO the number of the features are reduce. Finally the convolutional neural network is used to training the system and classifying the face.

The information base in this examination is completed from ORL and Yale data sets. The ORL (Olivetti Research Laboratory) face data set contains 400 pictures of 40 unique people, there are ten diverse grayscale pictures of every one of 40 unmistakable people. Pictures were caught at various occasions and they have various varieties including various articulations (open or shut eyes, grinning or not grinning) and facial subtleties (with or without glasses). Pictures were taken with a capacity to bear some shifting and turn of the face up to 20 degrees [9].

For assessing of the proposed technique, we utilized the mean squared mistake (MSE), Mean supreme rate blunder (MAPE) and R square strategy. The mean squared mistake (MSE) is appeared by:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \quad (4)$$

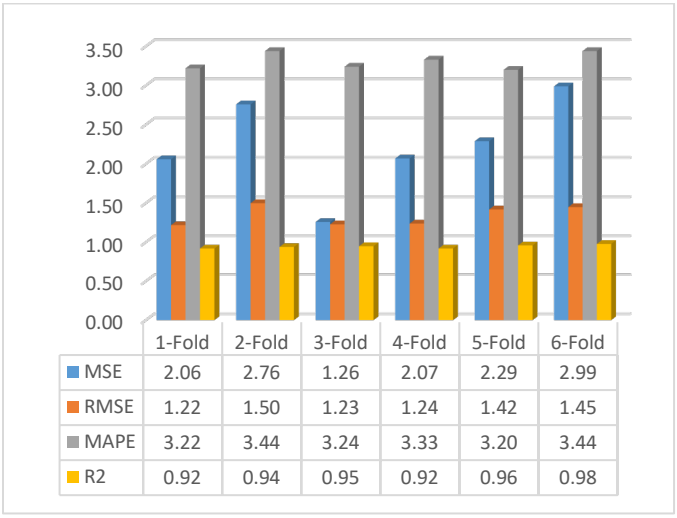
Mean absolute percentage error (MAPE) is shown in following equation:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|, \quad (5)$$

For R square we have the estimated value as:

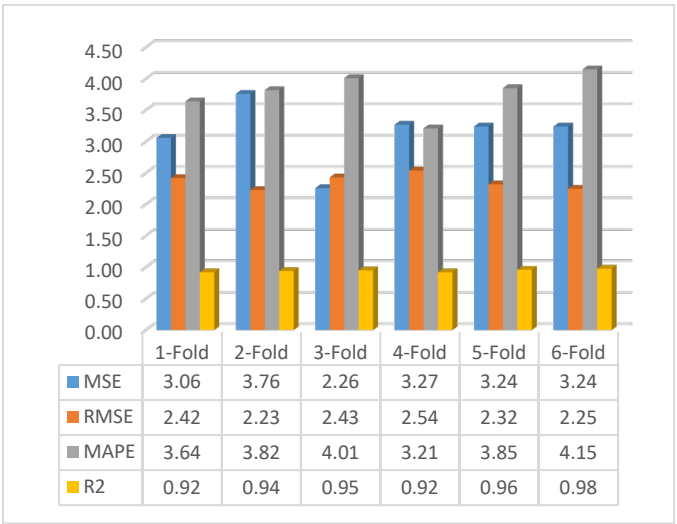
$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad (6)$$

The methodology was performed and tried utilizing genuine with balanced specimes from ORL and YALE datasets. The outcomes are appeared in figure 2 to figure 3. The aftereffects of testing the framework utilizing ORL dataset showed two things, first, how the preprocessing stage improve the exactness. Second, how might we union or breaker two techniques for include extraction to create an incredible third strategy that can achieve the work.



	1-Fold	2-Fold	3-Fold	4-Fold	5-Fold	6-Fold
MSE	2.06	2.76	1.26	2.07	2.29	2.99
RMSE	1.22	1.50	1.23	1.24	1.42	1.45
MAPE	3.22	3.44	3.24	3.33	3.20	3.44
R2	0.92	0.94	0.95	0.92	0.96	0.98

Figure 2. The MSE, RMSE, MAPE and R for train



	1-Fold	2-Fold	3-Fold	4-Fold	5-Fold	6-Fold
MSE	3.06	3.76	2.26	3.27	3.24	3.24
RMSE	2.42	2.23	2.43	2.54	2.32	2.25
MAPE	3.64	3.82	4.01	3.21	3.85	4.15
R2	0.92	0.94	0.95	0.92	0.96	0.98

Figure 3. The MSE, RMSE, MAPE and R for test

We have seen that the acknowledgment rate and exactness results from the investigations can't be fulfilling while using the Gabor wavelet and profound learning, due some variety of the estimations of highlights which prompts a defilement in the order step. Along these lines, when contrasted and Gabor wavelet includes, the assortment will be enormous consequently includes are between - 14 up to 254. Hence, ideal highlights should be chosen.

V. CONCLUSION

In this study, a method for face recognition is proposed. This method uses the Histograms of Oriented Gradients of the image to extract the features. The PSO method used to select the efficient features and convolutional neural network used to training the system and classification. Proposed technique utilize the balance methodology either in the picture space or in the element space, thus the element space is another approach to accomplish the advantage of utilizing the evenness property of the face to improve the acknowledgment rates. The proposed technique is inspected and tried for face acknowledgment utilizing information from ORL and YALE dataset and exploratory outcomes shows that it has a presentation higher than strategies in the writing.

ACKNOWLEDGMENT

The authors appreciate the Sefax University and Istanbul Ayvansaray University for supporting of this paper.

REFERENCES

[1] F. Zhao, J. Li, L. Zhang, Z. Li, and S.-G. Na, "Multi-view face recognition using deep neural networks," *Futur. Gener. Comput. Syst.*, vol. 111, pp. 375–380, 2020.

[2] S. Allagwail, O. S. Gedik, and J. Rahebi, "Face recognition with symmetrical face training samples based on local binary patterns and the Gabor filter," *Symmetry (Basel).*, vol. 11, no. 2, p. 157, 2019.

[3] U. Jayaraman, P. Gupta, S. Gupta, G. Arora, and K. Tiwari, "Recent development in face recognition," *Neurocomputing*, vol. 408, pp. 231–245, 2020.

[4] J.-Y. Gan and Y.-W. Zhang, "A new approach for face recognition based on singular value features and neural networks," *Acta Electron. Sin.*, vol. 32, no. 1, pp. 170–173, 2004.

[5] J. Gan, Y. Zhang, and S. Mao, "Adaptive Principal Components Extraction Algorithm and Its Application In the Feature Extraction of Human Face," *Acta Electron. Sin.*, vol. 30, no. 7, pp. 1013–1016, 2002.

[6] S. Ahmed, M. Frikha, T. D. H. Hussein, and J. Rahebi, "Optimum Feature Selection with Particle Swarm Optimization to Face Recognition System Using Gabor Wavelet Transform and Deep Learning," *Biomed Res. Int.*, vol. 2021, 2021.

[7] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, 2006.

[8] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Face recognition from a single image per person: A survey," *Pattern Recognit.*, vol. 39, no. 9, pp. 1725–1745, 2006.

[9] Z. Liu, J. Pu, Q. Wu, and X. Zhao, "Using the original and symmetrical face training samples to perform collaborative representation for face recognition," *Optik (Stuttg).*, vol. 127, no. 4, pp. 1900–1904, 2016.

[10] L. Liu, P. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Local binary features for texture classification: Taxonomy and experimental study," *Pattern Recognit.*, vol. 62, pp. 135–160, 2017.

[11] Y. Sun and J. Yu, "Facial expression recognition by fusing Gabor and local binary pattern features," in *International Conference on Multimedia Modeling*, 2017, pp. 209–220.

[12] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. image Process.*, vol. 19, no. 6, pp. 1635–1650, 2010.

[13] S. Anila and N. Devarajan, "Preprocessing technique for face recognition applications under varying illumination conditions," *Glob. J. Comput. Sci. Technol.*, 2012.

[14] M. M. Kasar, D. Bhattacharyya, and T. H. Kim, "Face recognition using neural network: a review," *Int. J. Secur. Its Appl.*, vol. 10, no. 3, pp. 81–100, 2016.

[15] X. Zou, J. Kittler, and K. Messer, "Illumination invariant face recognition: A survey," in *2007 first IEEE international conference on biometrics: theory, applications, and systems*, 2007, pp. 1–8.

[16] Y. Z. Goh, A. B. J. Teoh, and M. K. O. Goh, "Wavelet based illumination invariant preprocessing in face recognition," in *2008 Congress on Image and Signal Processing*, 2008, vol. 3, pp. 421–425.

[17] H. R. E. Doost and J. Rahebi, "An efficient method for texture classification with local binary pattern based on wavelet transformation," *Int. J. Eng. Sci. Technol.*, vol. 4, no. 12, pp. 4881–4885, 2012.

[18] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, 2005, vol. 1, pp. 886–893.

[19] F. P. Shah and V. Patel, "A review on feature selection and feature extraction for text classification," in *2016 international conference on wireless communications, signal processing and networking (WiSPNET)*, 2016, pp. 2264–2268.

[20] P. Yang and G. Yang, "Feature extraction using dual-tree complex wavelet transform and gray level co-occurrence matrix," *Neurocomputing*, vol. 197, pp. 212–220, 2016.

[21] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European conference on computer vision*, 2006, pp. 404–417.

[22] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[23] A. Eleyan and H. Demirel, *Pca and lda based neural networks for human face recognition*, vol. 558. INTECH Open Access Publisher, 2007.

[24] S. J. Dixon and R. G. Brereton, "Comparison of performance of five common classifiers represented as boundary methods: Euclidean distance to centroids, linear discriminant analysis, quadratic discriminant analysis, learning vector quantization and support vector machines, as dependent on," *Chemom. Intell. Lab. Syst.*, vol. 95, no. 1, pp. 1–17, 2009.

[25] G. Mehta and S. Vatta, "'An Introduction to a Face Recognition System using PCA, FLDA and Artificial Neural Networks," *IJARCSSE*, vol. 3, no. 5, 2013.

[26] S. Meshgini, A. Aghagolzadeh, and H. Seyedarabi, "Face recognition using Gabor-based direct linear discriminant analysis and support vector machine," *Comput. Electr. Eng.*, vol. 39, no. 3, pp. 727–745, 2013.

[27] O. Nikisins and M. Greitans, "Local binary patterns and neural network based technique for robust face detection and localization," in *2012 BIOSIG-Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG)*, 2012, pp. 1–6.

[28] C. Ravat and S. A. Solanki, "Survey on different methods to improve accuracy of the facial expression recognition using artificial neural networks," in *Proceedings of the National Conference on Advanced Research Trends in Information and Computing Technologies*, 2018, vol. 4, no. 2.

[29] C. Tomasi, "Histograms of oriented gradients," *Comput. Vis. Sampl.*, pp. 1–6, 2012.