

# The Using Of PCA, Wavelet and GLCM In Face Recognition System, A Comparative Study

Manar Abdulkaream Al-Abaji <sup>a</sup>      Meaad Mohammed Salih <sup>b</sup>

<sup>a,b</sup> Dept. of Computer Science, College of Education for Pure Science, University of Mousl, Iraq.

Manar198051@gmail.com      meaad\_muhammed@yahoo.com

Submission date:- 1/4/2018      Acceptance date:- 29/5/2018      Publication date:- 27/11/2018

---

**Keywords:** Face Recognition, PCA, DWT, GLCM, Euclidean distance.

---

## Abstract

The process of data dimension reduction plays an important role in any face recognition system because many of these data are repetitive and irrelevant and this cause a problem in applications of data mining and learning the machine. The main purpose is to improve the performance of recognition by eliminating repetitive features.

In this research, a number of data reduction techniques were used like: Principal Component Analysis, Gray-Level Co-occurrence Matrix and Discrete Wavelet Transform for extracting the most important features from the images of persons. A different number of training and testing images were used to compare the performance of each of the techniques above in the recognition process. Euclidean distance scale was used to get results.

## 1-Introduction

The distinction of biometrics refers to the automatic discrimination of individuals based on the vectors of the traits derived from their physiological or behavioral characteristics. The most commonly used biometric methods are face recognition[1]. Face recognition has become one of the most challenging tasks of distinguishing patterns in past decades. Face recognition applications include: surveillance, forensics and forensic applications, secure electronic banking, smart cards, to identify the individual in international transport centers, to control access and in many other areas[2][3].

The face is one of the oldest techniques used in people's recognition, it can be classified in to[3][4]:

- ❖ Structural method:- It relies on some face features like nose, eyes, eyebrows, mouth, chin, and relative relationships.
- ❖ Holistic method:- depends on the overall shape of the face (Global Shape of the Face).
- ❖ Hybrid method:-it relies on both holistic and Structural methods. It is generally used for 3D face recognition.

## 2-Dimension reduction

The main purpose of reducing dimensions is to improve the performance of the classification by removing duplicate features that are irrelevant[5], Dimension reduction can be performed in two ways, namely: selecting features and extracting features. The selection of features is the process in which a new set of features is not generated, but only a partial set of original features is selected, leading to a reduction in the reproduction. The extraction of features is the process of creating a new set of features, and the extraction of features can be a linear or nonlinear blending of the original features [5][6].

## 3-Principal Component Analysis (PCA)

PCA is a useful statistical method used for both extract features and reduce the dimension; It is Involves a mathematical procedure that transform a number of interrelated variables into a number of non-interrelated variables called Principle compounds[3]. These compounds are linked to the original variables by orthogonal transformation. This modulation is defined in such a way that the first orthogonal dimension has the greatest amount of variance among the images[7]. PCA divide face images into small groups of the features that characterize the images called Eigen faces which are reliable to represent both the face images in database and the new face images[8].

The main benefit of the PCA is that the Eigen Face method helps reduce the size of the database required to distinguish the face image involved in the testing process. The images used in the training are not stored directly, but the vectors are stored by dropping each of the images used in training to a set of Eigen Faces that have been calculated. The image of the face entered in the test is dropped to the Eigen Faces group itself, and its weight vector is computed and then compared to the vectors of the weights stored in the database for best match[3][8].

PCA method steps are as follows[3][8]:

1- calculate the average mean for M training images using the equation below:

$$average = \frac{1}{M} \sum_{n=1}^M training\ images(n) \quad \dots (1)$$

2- subtract the average mean from the training images using equation (2):

$$sub = training\ images - average \quad \dots (2)$$

3- Covariance Matrix will be computed using equation (3):

$$Covariance = \sum_{n=1}^M sub(n)sub^T(n) \quad \dots (3)$$

4- from Covariance Matrix, the Eigenvectors and Eigenvalues will be computed.

5- Sort the Eigenvalues and their corresponding Eigenvectors, take a subset of Eigenvectors with highest Eigenvalues, these subset called Eigen faces.

6- bring down the training images after subtract the average mean from them onto Eigen faces to get the features of each training sample.

#### 4-Discrete Wavelet Transform (DWT)

DWT is a widely used tool in image processing applications like image edge enhancement, compression, and feature extraction[6].

DWT is used to extract robust features in face recognition, It can degrade the images into primary building blocks that are well localized both in space and frequency[6]. DWT degrades the image into four sub bands (LL,LH,HL,HH), low frequency sub band (LL) contains maximum energy of the images, while high frequency sub bands (LH,HL,HH) contain the details of the image[5]. fig(1) shows configuration of 2-level wavelet decomposition.

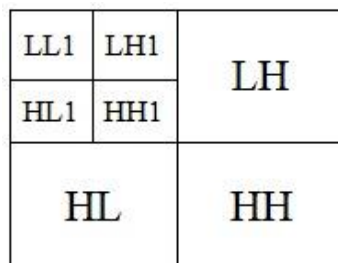


Fig.(1) configuration of 2-level wavelet decomposition.

#### 5-Gray-Level Co-occurrence Matrix

GLCM is a simple approach used to extract the textural features like Energy, Contrast, Correlation and Homogeneity from image. The using of Histogram in calculation carries only information related to density distribution; It does not convey any data for the spatial connections among pixels. The use of co-occurrence matrix which is a statistical method will help to extract features that describe the distribution of intensities and relative positions of neighboring pixels in an image[9][10].

If image H of dimension M×M, the co-occurrence matrix C can be written as[11]:

$$C(i, k) = \sum_{y=1}^M \sum_{z=1}^M \begin{cases} 1, & \text{if } H(y, z) = i \text{ and } H(y + \Delta y, z + \Delta z) = j \\ 0, & \text{other wise} \end{cases} \quad \dots (4)$$

Where ( $\Delta y, \Delta z$ ) indicates the distance between the pixel of interest and its neighbor, this makes the co-occurrence matrix sensitive to rotation, this can be avoided by choosing offsets sweeping through 180 degree at the same distance parameter  $\Delta$  to achieve a degree of rotational invariance (i.e.  $[0 \ \Delta]$  for  $0^\circ$  :P horizontal,  $[-\Delta, \Delta]$  for  $45^\circ$  :P right diagonal,  $[-\Delta \ 0]$  for  $90^\circ$  :P vertical, and  $[-\Delta, -\Delta]$  for  $135^\circ$  :P left diagonal)[11]. Fig (2) shows the co-occurrence matrix that calculated for each of the four different angles.

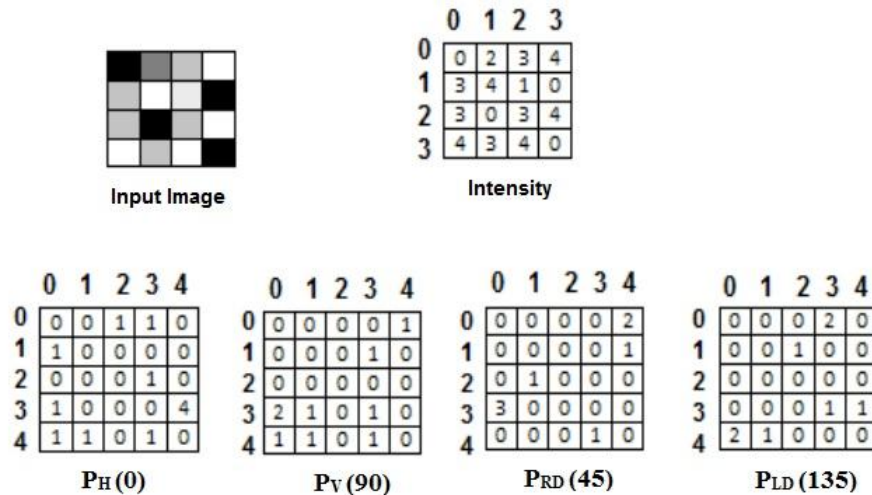


Fig.(2) co-occurrence matrix by using the four different angles, 0°, 90°, 45°, 135°

### 6-Distance Measurement

Once the features extraction method is applied on the images in database, the features vector is stored and then compared with the features vector of the test image using one of the distance measurement methods[5]. Measuring the distance between two image is the final step in face recognition. This distance describes the similarity between two image features vectors[8].

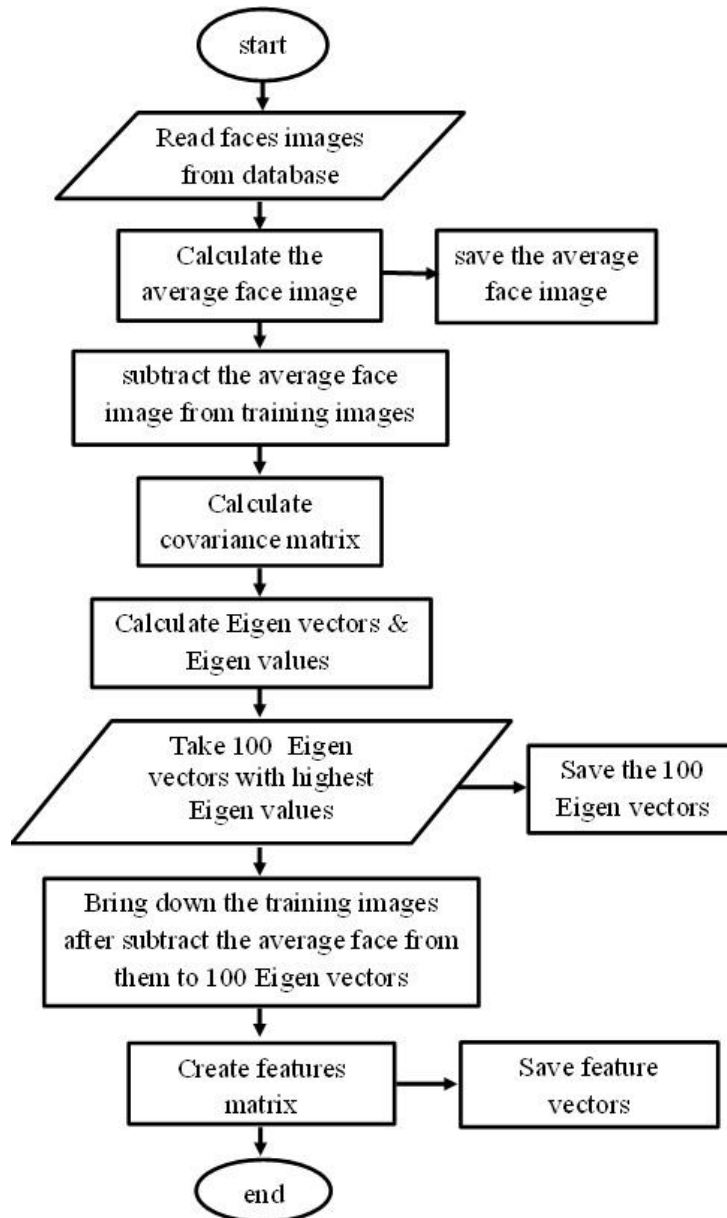
Euclidean distance is a typical and popular distance measurement method. It is defined as the straight-line distance between two points. Euclidean distance formula[8] :

$$\text{Euclidean distance}(X, Y) = \sqrt{\sum_{n=1}^{\text{No. of images}} (X_n - Y_n)^2} \quad \dots (5)$$

### 7-The Comparative Study

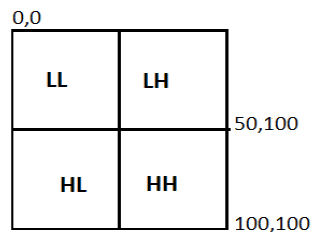
In any face recognition system, there are three main stages: the first stage gets face images, in this paper the database of face images have been configured by taking pictures of (10) people, each person has (40) images, the total number of face images collected is (400). A surveillance camera has been used to take pictures of people, Size of faces images is 100×100. The second stage extracts features from these image. Here, we used three features extraction methods :

1. PCA : each image in trained faces database has (10,000) value, to extract features from image Only (100) Eigen vectors have been adopted, and are also called Eigen faces with the highest Eigen values, which has highest contrast ratio between the face images used in the training, so we have reduced the graphic dimension of the image from (10,000) value to (100) value which will be entered to a stage of recognition .Fig.(3) shows the steps of PCA.



**Fig.(3) PCA method diagram**

2. Discrete Wavelet Transform : to reduces the image dimension(100 x100) , face image was been divided into four sections (LL, LH, HL, HH). The HH part of the image was taken only because it contains the detailed information of the image, HH part has been used as input to the recognition phase. The dimensions of this part (HH) are 50 x 50 as shown in Fig.(4), Fig.(5) shows the DWT steps.



**Fig.(4) (LL,LH,HL,HH) sections of image**

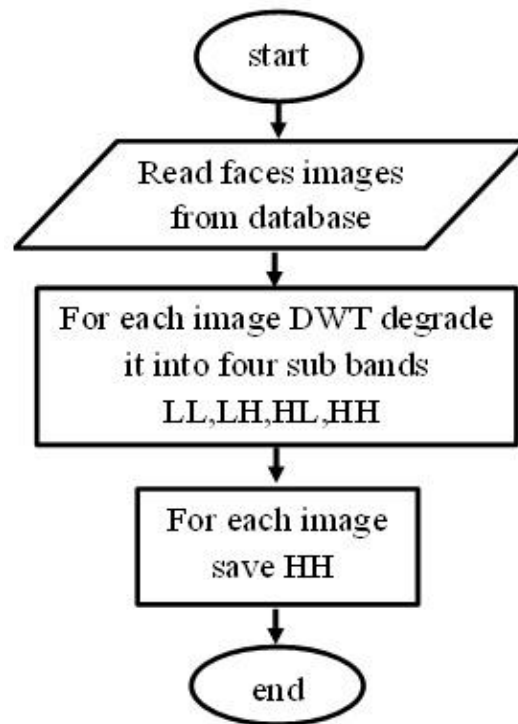


Fig.(5) DWT method diagram

3. GLCM: four image-related features: energy, contrast, correlation and homogeneity have been extracted using GLCM method. Four values for each feature according to the instruction and the following offset values. Total number of the values that were adopted in this method is (16) value which were used as input to discrimination, fig.(6) shows the GLCM steps.

[glcms ,SI]= graycomatrix (X, 'offset', offsets);

Offsets=[ 0 1 ; -1 1 ; -1 0 ; -1 -1]

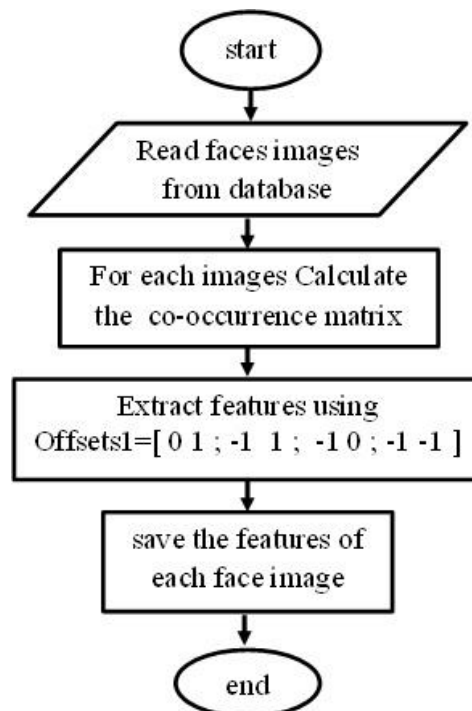


Fig.(6) GLCM method diagram

The final stage is the classification, we used Euclidean distance measurement to take the result of comparison, which calculates distances between the face image to be distinguished and all images used in the training, the image that has the least distance is the result of classification.

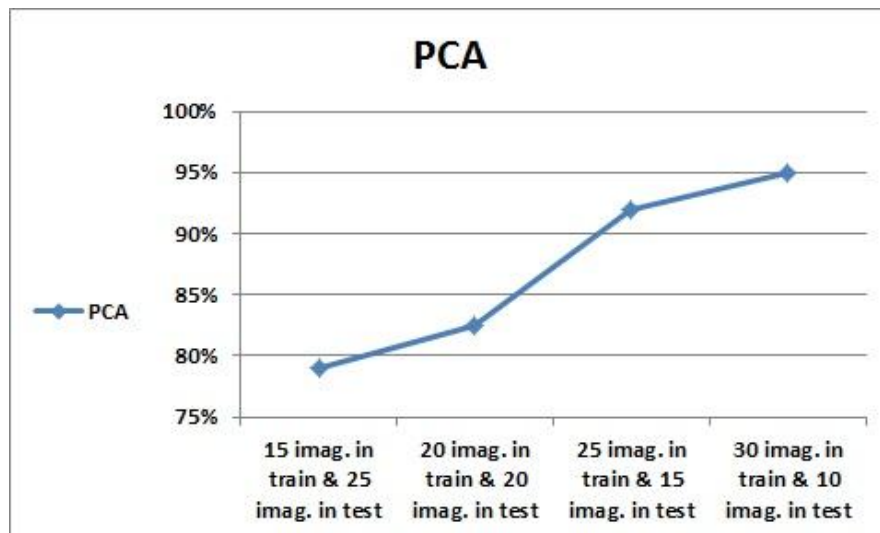
## 8-Results and Discussion

In the comparative study, experiment was conducted using gray images with same size 100×100 . simulations were done via using MATLAB.

When features were extracted from face images using PCA , a recognition rate of 79% was obtained through using (15) images in training and (25)images in testing, while the rate through using Wavelet is 82.4% and by using GLCM is 62% , then we used (20) images in training and (20) images in testing, we obtained recognition rate of 82.5% in PCA, while the rate obtained by using Wavelet is 85% and by using GLCM is 72% , this rate increased to 95% when we used PCA, 95% for Wavelet and 75% for GLCM when we used (30) images in training and (10)images in testing. Tables (1)(2)(3) and fig.(7)(8)(9) show the recognition rates when applying the three methods.

**Table (1) the recognition rates Using PCA**

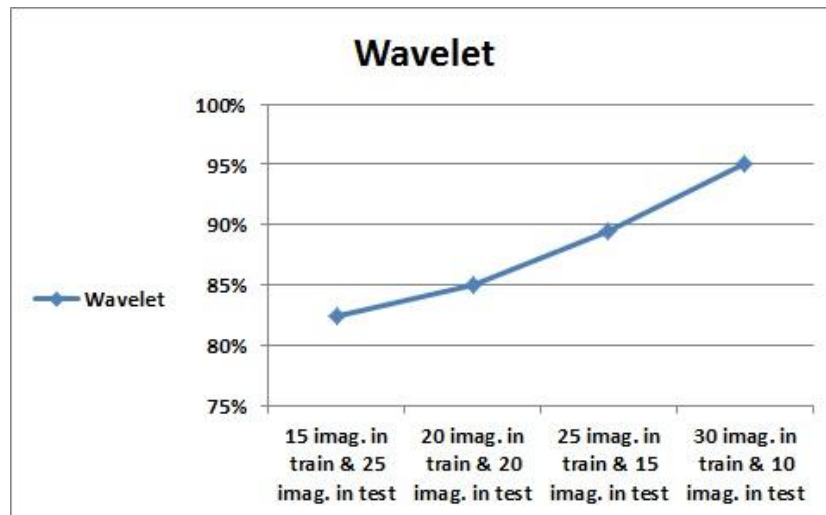
No.of images trained	No.of features	No.of images tested	Recognition rate
15	1500	25	79 %
20	2000	20	82.5 %
25	2500	15	92 %
30	3000	10	95 %



**Fig. (7) the recognition rates Using PCA**

**Table (2) the recognition rates Using Wavelet**

No.of images trained	No. of features	No.of images tested	Recognition rate
15	37,500	25	82.4 %
20	50,000	20	85 %
25	62,500	15	89.4 %
30	75,000	10	95 %



**Fig. (8) the recognition rates Using Wavelete**

**Table (3) the recognition rates Using GLCM**

No.of images trained	No.of features	No.of images tested	Recognition rate
15	240	25	62 %
20	320	20	72 %
25	400	15	74 %
30	480	10	75 %

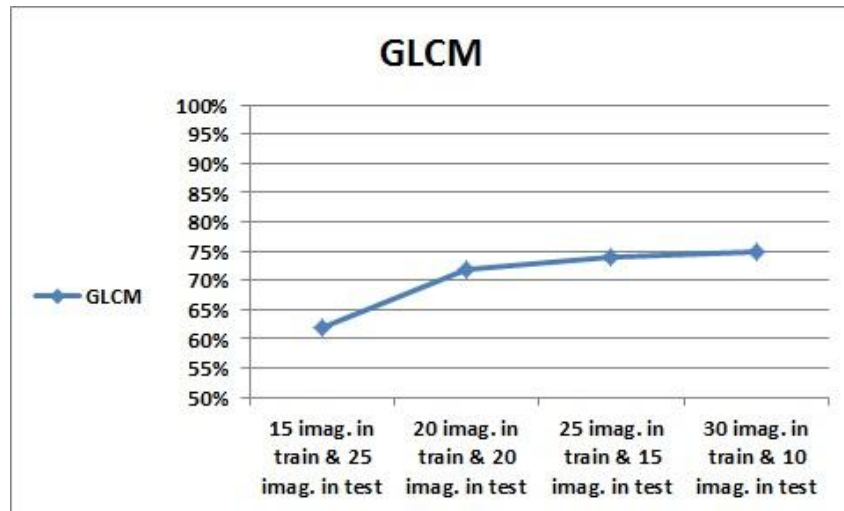


Fig. (9) the recognition rates Using GLCM

## 9- Conclusion

In this research:

- three methods were used to extract features from face images: PCA to extract (100) features, DWT to extract (50×50) features and GLCM to extract (16) features , and compare these methods to different numbers of training and test images for the purpose of knowing which method was better in the recognition process. The Euclidean distance scale was used to obtain the results.
- The results illustrate that, as the number of images used in the training increases and the number of images used in the test decreases, the rate of recognition increase.
- They also illustrate that the PCA and Wavelet are better than GLCM in Extract features from face images in face recognition system, fig.(10) shows the comparison among PCA, Wavelet and GLCM.

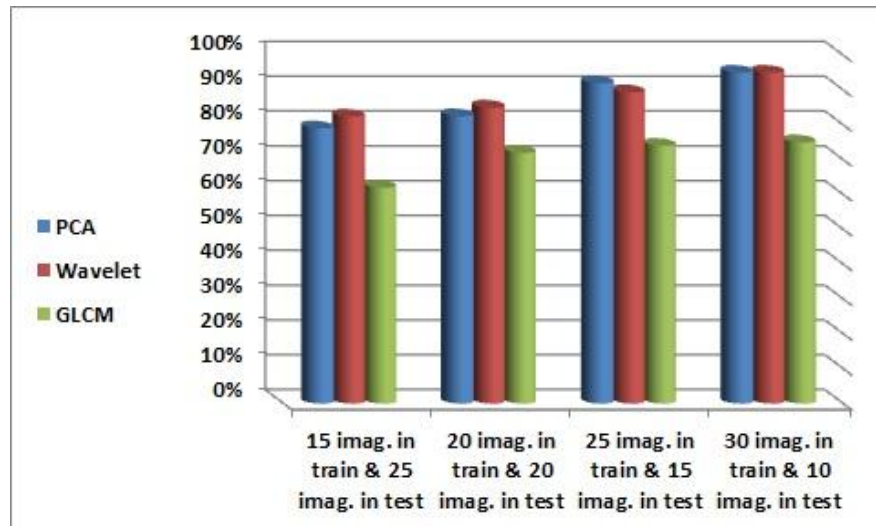


fig.(10) the comparison among PCA, Wavelet and GLCM.

## References

- [1] S. Dalali, "Daubechies Wavelet Based Face Recognition Using Modified LBP", *Procedia Computer Science* 93 (2016), PP. 344 – 350, 2016.
- [2] P. K. Harra and D. Aggarwal, "Hybrid Approach for Face Recognition Using DWT and LBP", *International Journal on Recent and Innovation Trends in Computing and Communication*, vol.5, Issue 7, pp.523-527, 2017.
- [3] N. H. Barnouti, "Face Recognition using PCA-BPNN with DCT Implemented on Face94 and Grimace Databases", *International Journal of Computer Applications* (0975 – 8887), Vol. 142, No.6, pp.8-13, 2016.
- [4] R. H. Pancholi and S. Swarndeep, "EFFICIENT FACE RECOGNITION METHOD FOR OCCLUDED IMAGES", *International Journal of Advance Research And Innovative Ideas In Education*, vol. 3, Issue 2, pp.5672-5679, 2017.
- [5] A. Ravindran, R. Aggarwal, N. Bhatia and A. Sharma, "Face Recognition using Evolutionary Approach", *International Research Journal of Computers and Electronics Engineering (IRJCEE)*, Vol. 3, Issue 1, 2015.
- [6] N. Adinarayana, B. A. Reddy and G. B. Reddy, "Comparative studies of feature extraction in face recognition system by using Haar wavelet transform", *International Conference on Advances In Computing ,Electrical and Communication Engineering*, Vol. 3, Issue.1, pp.60-63, 2017.
- [7] Taqdir and R.Dhir, "Face Recognition using SIFT Key with Optimal Features Selection Model", *International Journal of Advanced Computer Science and Applications (IJACSA)*, Vol. 8, No. 2, pp.403-409, 2017.
- [8] N. H. Barnouti, W. E. Matti, S. S. Al-Dabbagh and M. A. Naser, "Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance", *International Journal of Advanced Computer Science and Applications (IJACSA)*, Vol. 7, No. 5, pp.371-377, 2016.
- [9] A.Eleyanand H.Demirel, "Co-occurrence matrix and its statistical features as a new approach for face recognition", *Turkish Journal of Electrical Engineering and Computer Sciences*, vol.19, No.1, pp. 97-107, 2011.
- [10] S.Yamuna and S.Abirami, "Feature Extraction of Face Value Through Gray-Level Co-Occurrence Matrix", *International Journal of Open Information Technologies*, vol. 3, no. 6, pp. 32-35, 2015.
- [11] A. J. Dhanaseely, S. Himavathi and E. Srinivasan, "Performance Comparison of Neural Classifiers for Face Recognition System Using GLCM Features", *International Journal of Innovative Research in Computer Science & Technology (IJIRCST)*, Vol.4, Issue 1, pp. 15-18, 2016.

## الخلاصة

تلعب عملية التقليل البعدي للبيانات دوراً هاماً في أي نظام لتمييز الوجه، لأن العديد من هذه البيانات متكررة و غير ذي صلة وهذا يسبب مشكلة في تطبيقات التعلم الآلي والتتقيب عن البيانات. الغرض الرئيسي من التقليل البعدي للبيانات هو لتحسين أداء التمييز عن طريق إزالة الميزات المتكررة.

تم في هذا البحث استخدام عدد من تقنيات تقليل البيانات مثل تقنية تحليل المركبات الأساسية (PCA) و تحويل المويجة المنقطع (DWT) ومصفوفة الحدوث المشترك (GLCM). هدف البحث هو استخلاص الميزات الأكثر أهمية من الصور، تم استخدام اعداد مختلفة من الصور في التدريب والاختبار لغرض مقارنة الاداء لكل من التقنيات اعلاه في عملية التمييز، كما تم استخدام مقياس المسافة الاقليدية للحصول على النتائج.

**الكلمات الدالة:** تمييز الوجه، تحليل المركبات الأساسية، تحويل المويجة المنقطع، مصفوفة الحدوث المشترك، والمسافة الاقليدية.