Feature Fusion and Classifier Ensemble Technique for Robust Face Recognition

Hamayun A. Khan *h.khan@arabou.edu.kw*
Faculty of Computer Studies/ Arab Open University
Industrial Ardiya, 13033, Kuwait

Abstract

Face recognition is an important part of the broader biometric security systems research. In the past, researchers have explored either the Feature Space or the Classifier Space at a time to achieve efficient face recognition. In this work, both the Feature Space optimization as well as the Classifier Space optimization have been used to achieve improved results. The efficient technique of Feature Fusion in the Feature Space and Classifier Ensemble technique in the Classifier Space have been used to achieve robust and efficient face recognition. In the Feature Space, the Discrete Wavelet Transform (DWT) and the Histogram of Oriented Gradient (HOG) features have been extracted from face images and these have been used for classification purposes after Feature Fusion using the Principal Component Analysis (PCA). In the Classifier Space, a Classifier Ensemble has been used, utilizing the bagging technique for ensemble training, instead of a single classifier for efficient classification. Proper selections of various parameters of the DWT, HOG features and the Classification Ensemble have been considered to achieve optimum performance. The proposed classification technique has been applied to the AT&T (ORL) and Yale benchmark face recognition databases, and we have achieved excellent results of **99.78%** and **97.72%** classification accuracy respectively. The proposed Feature Fusion and Classifier Ensemble technique has been subjected to sensitivity analysis and it has been found to be robust under reduced spatial resolution conditions.

Keywords: Face Recognition, Feature Space, Wavelet Analysis, HOG Descriptors, Principal Component Analysis, Feature Fusion, Classifier Space, Classifier Ensemble, Linear Discriminant Classifiers, Robust Classification, Cross Validation.

1. INTRODUCTION

Biometric security systems are an important area of research and face recognition is a key element of the biometric identification systems [1, 2]. The recent proliferation of image capturing technologies and abundant availability of face images has given impetus to finding more efficient solutions to the face recognition problem. Face recognition is an essential part of the law enforcement work since Closed Caption TV (CCTV) systems and image capturing technologies are widely used for law enforcement purposes. Security systems can utilize efficient face recognition algorithms in order to identify any person wanted by law. Access control systems can use face recognition along with other access control technologies to enable access to premises or resources. Apart from law enforcement purposes, face recognition can be used in social networks, commercial and general web search applications. The availability of large amount of face images on the web makes face recognition important for web search applications. Popular social networking sites such as Instagram and Pinterest are heavily image focused. Searching large image databases on social networks can make use of face recognition techniques to perform broader and effective searches. The use of face recognition technology has to be balanced against the intended purpose and the various issues related to privacy and confidentiality. The need to apply face recognition technology in order to identify and catch, say, a purse snatching thief is understandable on one hand but to apply large scale face recognition systems to tag photos on the other hand may involve privacy issues which need to be addressed carefully [1, 2].

Past research in the area of Face Recognition has focused either to exploit the Feature Space or the Classifier Space at a time for face recognition. In this work a new methodology for face recognition has been proposed to optimize processes both in the Feature Space as well as in the Classifier Space for achieving efficient face recognition. In the Feature Space, it is proposed to extract 3 sets of features from face images and then to use Feature Fusion to fuse them together through PCA technique to get a compact feature vector. In the Classifier Space it is proposed to use a Classifier Ensemble instead of a single classifier for classification purposes. By optimizing the processes both in the Feature Space as well as in the Classifier Space, it will be shown that superior classification performance is achievable. Also, it will be shown that the proposed technique is robust under degradation of spatial resolution conditions since the classification performance of the proposed system is slightly degraded when the spatial resolution is reduced by half.

The breakdown of this article is given below. A review of previous work on face recognition is presented in Section 2 of this paper. In Section 3, the proposed robust system for face recognition has been introduced. The Feature Space and the Classifier Space optimizations have been discussed in detail in Section 4. The overall setup of the proposed system has been presented in Section 5. Results and discussions have been presented in Section 6. Comparison of results with existing techniques has been discussed in Section 7. Section 8 presents the conclusions and future work ideas.

2. RELATED WORK

The problem of face recognition has been studied extensively for a number of years. Researchers have applied a wide variety of techniques to explore the face recognition problem. Excellent surveys on face recognition systems and techniques are available in [3, 4]. Chihaoui et al. [3] have provided a comprehensive survey of face recognition systems till the year 2016, whereas Jafri and Arabnia [4] have covered the research period till the year 2009. As discussed in great detail in [3], the 2-Dimensional (2D) face recognition techniques can be broadly divided into 4 categories i.e. Global (or Holistic) Approaches, Local Approaches, Hybrid Approaches, also including Statistical Models, and Modern Techniques. In the first category of the Global Methods, the whole face image is considered globally without the need to extract regions of interests such as eyes, mouth, nose, etc. The techniques in this category are further subdivided into Linear and Non-Linear methods. In the Linear techniques, the images are mapped into a smaller sub-space using linear projections [3]. The Linear techniques include the classical Eigenface Methodology [5--7], 2D Principal Component Analysis (PCA) [8], Independent Component Analysis (ICA) [9], Linear Discriminant Analysis (LDA) [10], etc. In the Non-Linear methodology, a ‘kernel’ function is used to transform the problem into another larger space in which the problem becomes linear. The Non-Linear techniques include the Kernel PCA (KPCA) [11], Kernel Direct Discriminant Analysis [12], Kernel ICA (KICA) [13], ISOMAPs [14], etc.

In the second category of Local Approaches, some particular facial features are used for recognition purposes [3]. The various techniques in this category include the Dynamic Link Architecture (DLA) based on the deformable topological graph [15], Elastic Bunch Grapg Matching (EBGM) [16], Geometric Feature Vector [17], Gabor Filters [18], Scale-Invariant Feature Transform (SIFT) [19], Local Binary Patterns(LBP) [20], etc.

In the third category of Hybrid Approaches, a combination of the Global and Local techniques is used. This category also makes use of Statistical Models i.e. a statistical model representing the face is constructed and used for recognition purposes [3]. The various techniques in this category include Hidden Markov Models (HMMs) [21], the Gabor Wavelet Transform based on the Psuedo HMM (GWT-PHMM) [22], Discrete Cosine Transform HMM (DCT-HMM) [23], HMM-LBP [24], etc. The fourth category of Modern Techniques [3] include 3D face recognition [25--27], Multimodal systems [28], Infrared systems [29--31] and Deep Learning Sytems [32--36]. The Deep Learning Systems are currently quite popular for Machine Learning and face recognition tasks and they generally use Convolutional Neural Networks (CNNs) for image recognition requiring large datasets for learning purposes. Big Corporations such as Google and Facebook have their own Deep Learning based systems with high accuracies including Deepface and Facenet respectively [37, 38]. Opensource software libraries, such as OpenFace, are also available for face recognition purposes [39].

In this work, a new methodology for face recognition has been proposed which differentiates itself from current technologies in that it optimizes both the Feature Space as well as the Classifier Space instead of focusing on one at a time. An initial version of this work using only the DWT based feature set was presented by the authors in [40]. A new feature set based on the HOG descriptors has been used in the current work along with DWT features and this has enabled us to achieve improved classification accuracy. The current work has also subjected the system to detailed sensitivity analysis using lower spatial resolution images. Three different feature sets are used in the Feature Space and these are fused together to achieve an optimized feature set that maximizes classification accuracy. In the Classifier Space, the Classifier Ensemble technique is used instead of a single classifier to achieve improved classification results. These optimizations in both the spaces have enabled us to develop a robust and efficient face recognition system.

3. THE PROPOSED ROBUST FACE RECOGNITION SYSTEM

In this work, a robust face recognition system is proposed based on optimizing the various processes both in the Feature Space as well as in the Classifier Space. The system architecture of the proposed face recognition system is shown in Figure 1 below.

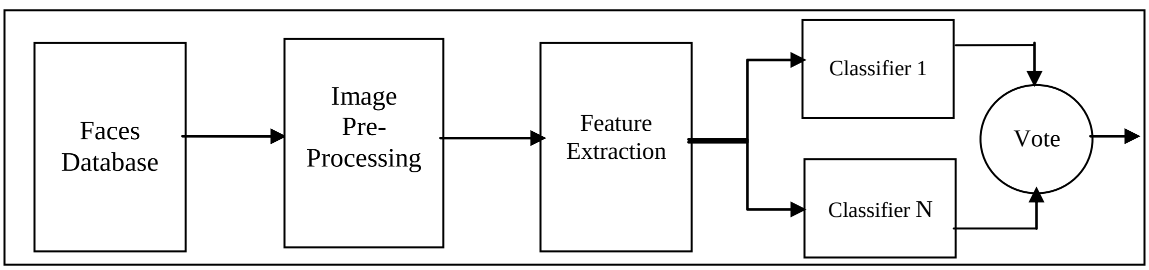

FIGURE 1: Proposed Robust Face Recognition System.

The proposed system mainly consists of the Feature Space processing and then the subsequent Classifier Space processing. Face Detection is also an essential part of face recognition systems but since the benchmark face databases that are considered in this work include mainly frontal images therefore the Face Detection part has not been used. The Feature Space processing includes the Pre-processing step, the Feature Extraction step and the subsequent Feature Fusion step. For Feature Extraction the 2-D DWT decomposition and the HOG descriptors have been used. Feature Fusion is based on the PCA and it is used to fuse the extracted features into a compact feature vector.

The Classifier Space processing consists of the Classifier Ensemble which is constructed from a number of individual base classifiers. In this work the Classification Ensemble class of Matlab with Linear Discriminant Analysis (LDA) classifiers as base classifiers have been used to construct the ensemble system. A set of 30 base classifiers constitute the ensemble. The Bagging Technique has been used to train the ensemble for face recognition task and the 10-fold cross validation technique has been employed to validate the classifier system. 50 runs of our simulations are performed and then the average of these 50 runs is taken as the final performance of our system. The classification performance is presented in the form of classification accuracy. The proposed system is based on the concept of optimising the various processes both in the Feature Space as well as in the Classifier Space. Details about the proposed recognition system are given in the following section.

4. FEATURE SPACE AND CLASSIFIER SPACE OPTIMIZATION

4.1 Benchmark Face Databases

Two benchmark face image databases i.e. the AT&T Database of faces [41], formerly known as the Olevetti Research Lab (ORL) Database of faces, and the Yale face image database [42] have been used for experimental purposes. Images of two spatial resolutions i.e. size 64X64 pixels and 32X32 pixels for both databases have been considered in this work.

The AT&T (ORL) database of face images is maintained by AT&T laboratories, Cambridge and it consists of images of 40 subjects each having 10 face images in the set for a total of 400 images [41]. The face images in the database are mostly frontal images with moderate facial expressions. The database does not suffer from any major Pose, Illumination and Expression (PIE) problem but it has slight Illumination and Facial Expression variations. The original face images in this database are of size 92x112 but images of two reduced spatial resolutions i.e. 64X64 pixels and 32X32 pixels have been considered in this work. A snapshot of images from the AT&T (ORL) Database is shown in Figure 2. The main reason for selecting smaller sizes for images is to evaluate the performance of the system on reduced spatial resolution images.

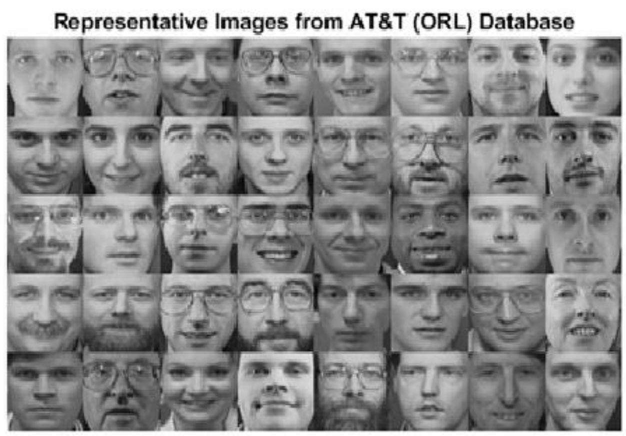

FIGURE 2: A snapshot of the images from AT&T (ORL) Database.

The second benchmark face database considered in this work is the Yale Database [42]. This database consists of face images of 15 subjects with 11 images each for a total of 165 images. This database is more challenging than the AT&T (ORL) Database since it exhibits comparatively more Illumination and Expression variations. The original face images in this database are of sizes 320x243 and 100x80 but here too images of 64X64 and 32X32 spatial resolution have been considered in this work for the same reason as mentioned above for the AT&T (ORL) Database. The various facial expressions exhibited by the subjects in this database include several expressions, such as smiling faces, yawning, winking and frowning. A snapshot of images from the Yale Database is shown in Figure 3.

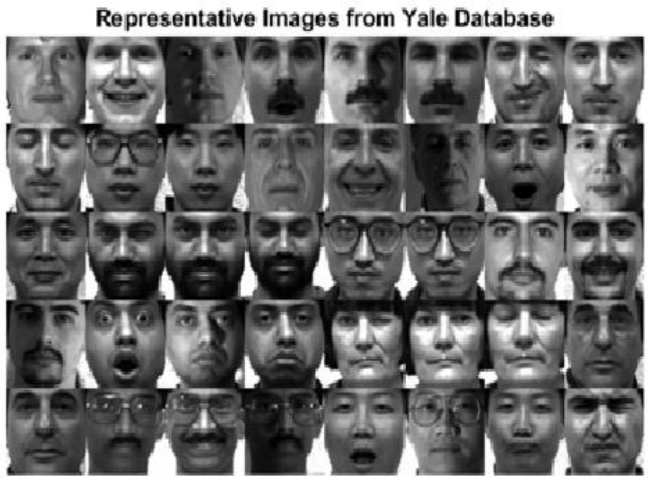

FIGURE 3: A snapshot of the images from Yale Database.

4.2 Image Pre-Processing Step

The image pre-processing step is usually required for the images that are challenging to classify. Images that exhibit the Pose, Illumination and Expression (PIE) problem fall in the challenging category of images for classification tasks. The partial PIE problem is exhibited by the Yale database compared to the AT&T (ORL) database. In this work, the adaptive histogram equalization technique has been used in order to alleviate the illumination problem from images and make them more suitable for classification. An improvement in classification accuracy has been achieved when the pre-processing is applied to the Yale database. Preprocessing is not usually required for the AT&T (ORL) database since the images are mostly frontal with little PIE problems.

4.3 HOG Based Features Extraction

The Histogram of Oriented Gradient (HOG) is an efficient technique that has been applied for human detection [43]. The performance of HOG based classification depends on the proper selection of cell size used for processing the face images to extract the HOG descriptors. This technique involves proper selection of the 2 element vector for the cell sizes. Figure 4 below depicts this phenomenon. We have observed that selecting a very small cell size, say 2x2 pixels, compared to face image of size 64X64, results in a large size HOG feature vector that tends to over burden the classifiers resulting in lower classification accuracy. On the other hand, selecting a very large cell size, say 10x10 pixels, results in a feature vector of smaller size which again is inappropriate for classification purposes. It has been experienced in this work that cell sizes of 7x7 pixels and 5X5 pixels for face images of size 64X64 and 32X32 respectively result in optimum classification performance.

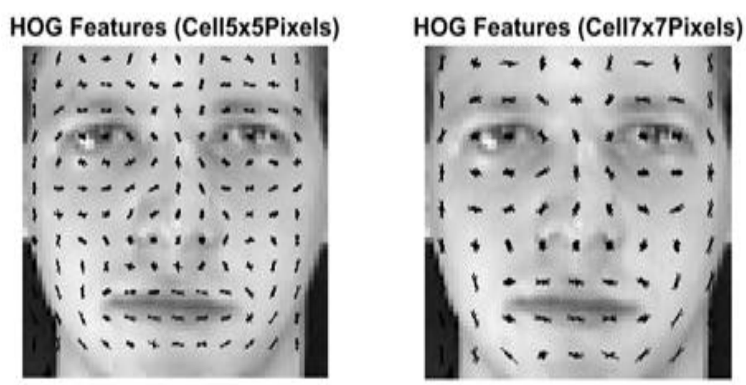

FIGURE 4: Effect of Cell Size for HOG Feature Extraction.

Hence for the HOG based feature set, we have used cell sizes of 7x7 pixels and 5x5 pixels for the 64x64 and 32x32 spatial resolution face images respectively. The resulting feature vector based on HOG descriptors is denoted by 'Features_HOG'. As an example, for a 64X64 size image, we use a size 7x7 pixels cell element to compute the HOG features. This results in a 2304x1 size HOG feature vector.

4.4 DWT Based Feature Extraction

The Discrete Wavelet Transform (DWT) is a well-established technique for image analysis [44, 45]. It enables multi-resolution analysis of images at different scales and resolutions. Thus it is possible to extract rich wavelet based features from images which can be used for classification purposes. The DWT can be expressed by the following set of equations [45]:

$$W_{\varphi}(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f\ (x, y)\varphi_{j_0,m,n}(x, y) \tag{1}$$

$$W_{\Psi}^{i}\ (j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f\ (x, y)\ \Psi_{j,m,n}^{i}\ (x, y), \tag{2}$$

$$i = \{H, V, D\}$$

The functions φ (x,y) and ψ (x,y) are the scaling function and wavelets respectively. We perform 2 levels of decomposition of the face images and the computed coefficients form the DWT feature vector for each image. The DWT decomposition results in both the Detailed and the Approximation Coefficients. We use the Approximation Coefficients as feature vectors since they carry useful information about the images. A Level 2 DWT decomposition of one of the face images from the AT&T (ORL) database is shown in Figure 5 below. Hence we have used Level 1 and 2 DWT decompositions in our work to extract optimum DWT features, denoted as 'Approx_Coeffs_DWT_Level1' and 'Approx_Coeffs_DWT_Level2', for improved classification accuracy.

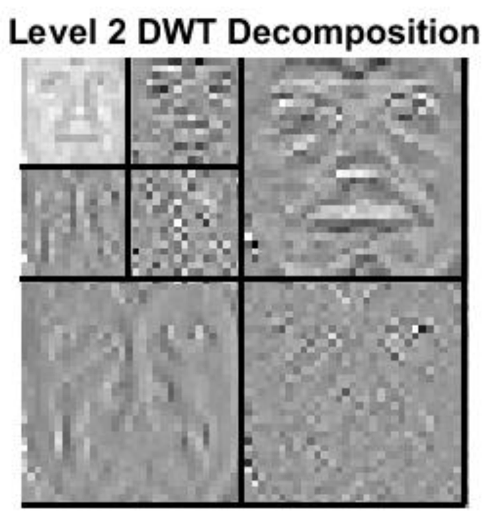

FIGURE 5: Wavelet Analysis of a Sample Face Image.

For the individual Wavelet Function, the Daubachies Wavelet ‘db1’ has been used. For instance, for a size 64X64 image, we compute the approximation coefficients from the image at 2 decomposition levels i.e Level1 and Level2. We use the Approximation Coefficients as feature vectors based on DWT coefficients resulting in a 1024X1 DWT feature vector at Level1 and 256x1 at Level2. Thus the DWT based feature vector consists of the following:

Features_DWT=[Approx_Coeffs_DWT_Level1 Approx_Coeffs_DWT_Level2] (3)

4.5 Feature Fusion for Robust Face Recognition

The individual feature vectors computed using the HOG and DWT analysis i.e. Features_HOG and Features_DWT are normalized before concatenating them to form the overall feature vector. The normalization is essential so as to ensure that individual feature vector values do not overshadow other values. After normalization, the DWT and HOG features of each image are put together to form an overall feature vector.

Features_Overall = [Features_HOG Features_DWT] (4)

The 2 level decompositions of the images by the DWT analysis lead to high dimensional feature vectors. Also, the HOG analysis results in high dimensional feature vector. Using high dimensional feature vectors for classification purposes often leads to poor performance of classifiers. Hence it is desirable to reduce the dimensionality of feature vectors to reduce computational cost as well as to achieve better performance. We utilize the PCA technique for dimensionality reduction as well as the fusion of the individual feature vectors to achieve the final feature vector, denoted as 'Features_Fused', which is used for classification purposes.

Features_Fused = PCA{ Features_Overall } (5)

The fusion of feature values results in better performance of the ensemble classifier since it utilizes the combined information contained in the DWT and the HOG analysis. Choosing an appropriate dimension for Feature Fusion has been based on the resulting best performance of the classification system. For instance, for a 64X64 size image AT&T(ORL) image, we have used a PCA Dimensionality of 50 for Feature fusion since it maximizes the classification accuracy.

The sample feature vector values for the AT&T (ORL) and Yale databases obtained after fusion, using dimensionality of 50, are shown in Figure 6 below. The figure shows plots of values for three different classes for each database. It is clear from the figure that same class values show strong similarity trends whereas the intra-class differences are also evident thus highlighting the distinguishing characteristics of the features.

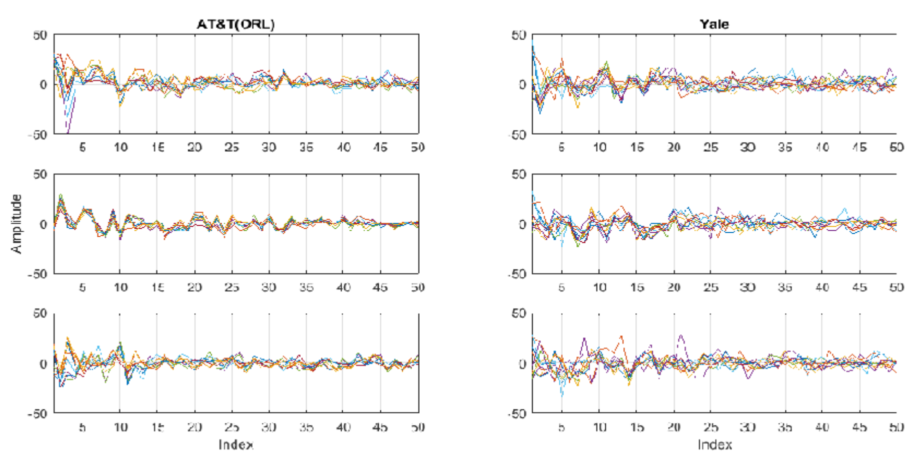

FIGURE 6: Sample Fused Feature values (Dim=50) for the 2 Databases for 3 classes.

4.6 Classifier Space Optimization Through LDA Based Ensemble

At the classification stage, an Ensemble of Classifiers has been used instead of a single classifier. The Ensemble or team of classifiers approach is superior to using single classifier since it uses the combined or collective classification capabilities of the ensemble to achieve superior classification performance. As mentioned above, the ensemble system used in our work uses 30 Linear Discriminant Analysis (LDA) based classifiers as base classifiers. The base classifiers are usually weak classifiers but by acting together in the ensemble system as a team the overall performance of the ensemble becomes superior to the base classifiers. The choice of LDA base classifiers is important since these maximize the between-class distances and minimize the in-class separations thus leading to higher classification accuracies.

The ensemble systems are commonly trained using the Boosting or the Bagging methodologies for ensemble learning [46]. In this work the ensemble has been trained on the face image features data obtained after feature fusion using the Bagging Technique for ensemble learning. Bagging or “bootstrap aggregation” uses resampled version of features data to train the ensemble. The cross validation methodology has been used to assess the performance of the ensemble system on the 2 benchmark databases. This methodology is useful in situations where the data set is not large enough to partition it into an independent Training Set and Test Set. Specifically, we have used the 10-fold cross validation methodology in our work to judge the classification accuracy of our system.

5. OVERALL SYSTEM SETUP

The overall setup of the proposed face recognition system depends on the architecture of the system as shown in Figure 2. The setup information of our system is related to the sub-blocks of our system and this information is presented in Table 1 below. For the Face Database sub-block, we have considered the Yale and AT&T (ORL) benchmark databases in our work. We have applied preprocessing operation only to the Yale database images as they suffer from the PIE problem. For feature extraction purposes, we have used the DWT and the HOG analysis based techniques. We have used 2 decomposition levels for our DWT based feature set and we have used the Wavelet db1 in our computations. For the fusion sub-block, the PCA technique has been used for dimensionality reduction and to fuse together the individual feature vectors to achieve an efficient feature vector that packs the discriminating powers of the individual features. For the classification part, we have used an ensemble of 30 Linear Discriminant Analysis (LDA) based classifiers as base classifiers in our classifier ensemble.

Item/Factor	AT&T (ORL) Database	Yale Database
Total subjects in Database	40	15
Number of images of each subject	10	11
Total Number of Images	400	165
Image resolutions	Size 64X64 & 32X32	Size 64X64 & 32X32
Preprocessing of images	None	Adaptive Histogram Equalization
Feature extraction methodology	2-D DWT and HOG Features	2-D DWT and HOG Features
Wavelets used	Daubechies "db1"	Daubechies "db1"
Wavelet decomposition levels	Level 1 & 2	Level 1 & 2
HOG cell size for 64x64 images	7x7 pixels	7x7 pixels
HOG cell size for 32x32 images	5x5 pixels	5x5 pixels
Feature Fusion Technique and Dimensionality reduction	PCA	PCA
Type of Base Classifiers	Linear Discriminant	Linear Discriminant
Number of base classifiers	30	30
Testing/Validation methodologies	10-Fold Cross Validation	10-Fold Cross Validation
Number simulation runs	50	50

TABLE 1: The Overall System Setup Information.

6. RESULTS AND DISCUSSION

A number of experiments have been performed to evaluate the performance of the face recognition system which is based on Feature Fusion and Classifier Ensemble. Sensitivity analysis has also been performed to evaluate the performance of the system on reduced spatial resolution images. The results obtained on the two selected benchmark databases are discussed below.

6.1 Performance on the AT&T (ORL) Face Database

The performance of our proposed face recognition system on the AT&T (ORL) Face Database, size 64X64, is shown in Figure 7. The figure shows that the classification accuracy of the system improves as the number of base classifiers increases in the ensemble. The best performance of 99.78% classification accuracy has been achieved for feature fusion vector using PCA Dimensionality of 50.

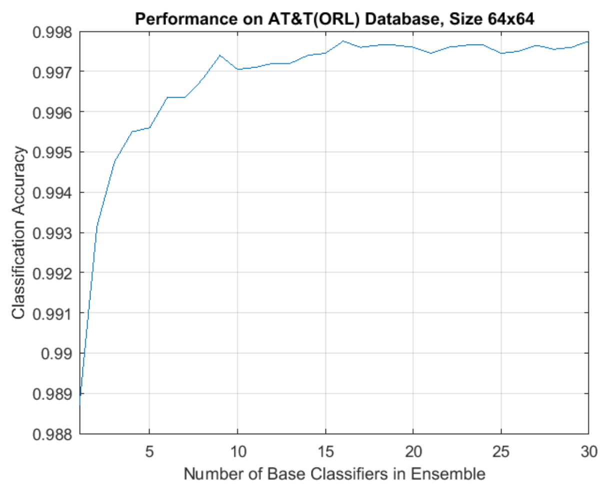

FIGURE 7: Performance on the AT&T (ORL) Database of size 64X64.

6.2 Performance on the Yale Face Database

The performance of our face recognition system on the Yale Face Database, size 64X64, is shown in Figure 8. Once again, the performance of the classification system improves as the number of base classifiers is increased in the ensemble system. However, compared to the AT&T (ORL) Face Database, the performance of this database is slightly less with a 97.72% classification accuracy. The PCA Dimensionality used for feature fusion is 50 in this case.

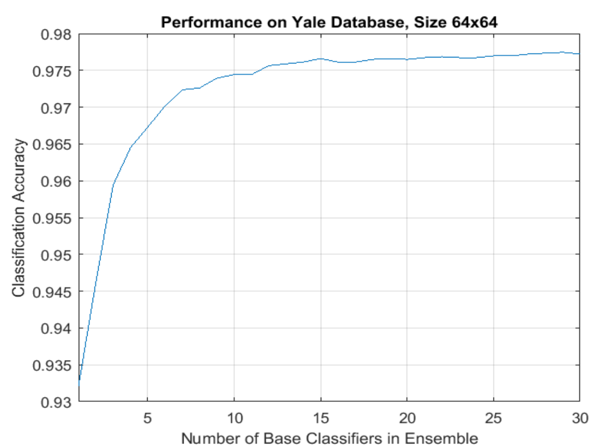

FIGURE 8: Performance on the Yale Database of size 64X64.

6.3 Sensitivity Analysis

We have further evaluated the performance of our Face Recognition System using face images of lower spatial resolution than the ones discussed above. The results for AT&T (ORL) and Yale databases at lower spatial resolution of 32X32 instead of 64X64 are shown in Figure 9 and Figure 10 respectively. A PCA dimensionality of size Dim=35 has been used for the low size images. We have obtained accuracies of 99.05% and 96.21% for the AT&T (ORL) and the Yale databases respectively. The results indicate a slight reduction in performance of the system for lower spatial resolution images.

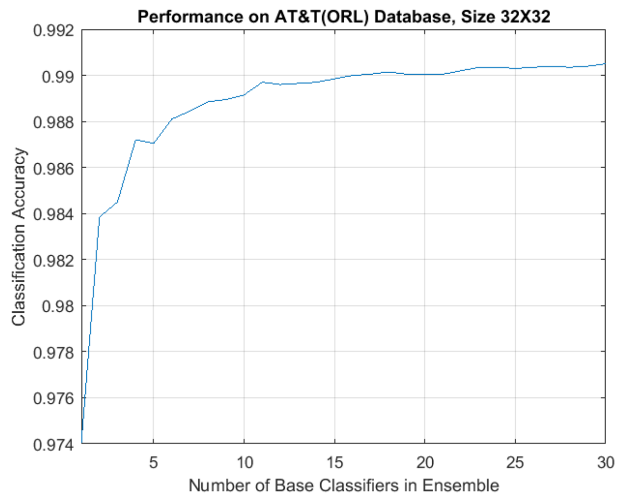

FIGURE 9: Performance on the AT&T (ORL) Database of size 32X32.

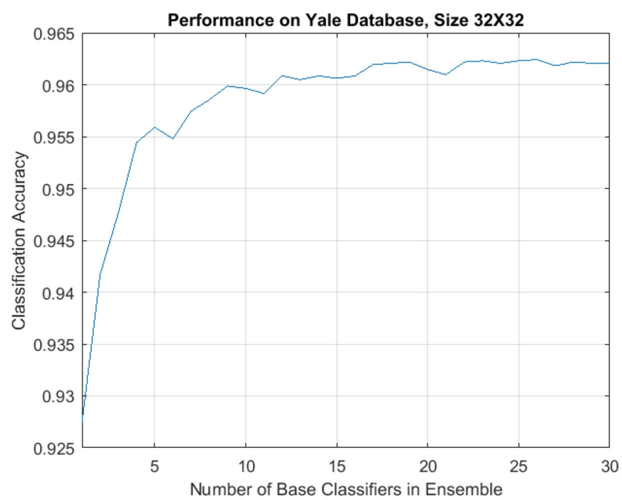

FIGURE 10: Performance on the Yale Database of size 32X32.

6.4 Summary of Results

The summary of results of our experiments and simulations is presented in Table 2 below.

Benchmark Database	Image Resolution	PCA Dimensions	Classification Accuracy Achieved
AT&T (ORL)	64X64	50	99.78%
Yale	64X64	50	97.72%
AT&T (ORL)	32X32	35	99.05%
Yale	32X32	35	96.21%

TABLE 2: Summary of Experimental Results.

It is clear from the table that highest classification accuracy of 99.78% has been achieved for the AT&T (ORL) database of size 64X64 pixels. For the Yale database of similar spatial resolution, we have obtained a classification accuracy of 97.72%. The better performance of our classification system on the AT&T (ORL) database compared to the Yale one is due to the challenging nature of the Yale database as it exhibits the PIE problem. The sensitivity analysis of the system shows that our system is robust against any spatial resolution degradation as it experiences a slight reduction in performance for lower resolution images of both databases.

7. COMPARISON OF RESULTS WITH STATE-OF-THE-ART TECHNIQUES

For the ORL database, researchers in the past have obtained classification rates ranging from 75.2% to 100% by employing various recognition techniques as discussed in Section 2 of this paper. The 99.78% classification performance of our system for the AT&T (ORL) images of size 64x64 pixels is excellent compared to existing techniques since our system has been subjected to an extreme validation strategy of 10-fold cross-validation along with the rigor of repeating the experiments a number of times and then selecting the average classification accuracy of the runs. This rigorous process highlights the robustness of our proposed technique along with the generalization property of our system to unseen data. Our proposed system is thus superior to any other system, which may have obtained a higher classification accuracy on a select few folds without any substantial validation through additional simulation runs. Our system is also capable of achieving 100% accuracy for some of the selected folds but we decided to determine a detailed and comprehensive performance through extensive experimentation rather than selecting a preferred fold with 100% accuracy. Also, the 64x64 size images used in our system compared to the larger 112x92 sizes of the original AT&T (ORL) database represent the higher robustness of our system compared to existing techniques that have been validated on higher spatial resolution images. For the Yale database, the classification performances achieved by previous researchers range from 84.24% to 99.5% accuracies. In this case too, the classification accuracy of 97.72% achieved by our system is excellent taking into account the robust nature of our system achieving this accuracy for reduced spatial resolution images of 64x64 size compared to 320x243 size images in original Yale database.

Our proposed Feature Fusion and Classifier Ensemble technique is a valuable contribution to the field of face recognition and it will have positive impact on the developments in this field. Comparison of our results with the State-of-the-Art techniques for face recognition clearly highlights the importance of the proposed technique. Simulations performed in this work have achieved classification accuracies of 99.78% and 97.72% on the AT&T (ORL) and Yale benchmark databases respectively, representing valuable contributions to the research effort in this field.

8. CONCLUSIONS AND FUTURE WORK

In this paper, we have proposed a robust face recognition system based on Feature Fusion and Classifier Ensemble techniques. We have obtained classification accuracy of 99.78% for the AT&T (ORL) Database and 97.72% for the Yale Database using images of 64X64 spatial resolution. Classification accuracy of 99.05% and 96.21% have been achieved for reduced resolution AT&T (ORL) and Yale Databases respectively. Our face recognition system uses Feature Fusion and an LDA based classification ensemble of 30 Base classifiers and 10-fold cross validation. Feature vectors have been extracted from face images using the DWT decomposition and HOG analysis. Dimensionality reduction and Feature Fusion have been achieved using the PCA technique. The results achieved in this work have demonstrated the effectiveness of our proposed system. It has been shown that using Feature Space Optimization through the Feature Fusion Technique and Classifier Space Optimization through Classifier Ensemble Technique has resulted in better classification performance instead of using concatenated high dimensionality feature vectors and single classifier.

The performance of our classification system is influenced by a number of factors such as the type of base classifiers, the number of base classifiers, the ensemble learning technique, type of features, dimensionality of feature space etc. As future research directions, we plan to consider various combinations of these factors in our research. We plan to consider other face image databases, such as the Extended Yale database, and other learning schemes for the ensemble system training. Also, future work would consider other types of base classifiers as members of the ensemble.

9. ACKNOWLEDGEMENTS

This work would not have been possible without the great research support provided by the Arab Open University (AOU) and the author greatly acknowledges this support.

10. REFERENCES

[1] A. K. Jain, A. Ross, and K. Nandakumar. *Introduction to Biometrics*. US, Springer, 2011.

[2] R. Jiang, S. Al-Madeed, A. Bouridane, D. Crookes, and A. Beghdadi, eds. Biometric Security and Privacy: Opportunities & Challenges in The Big Data Era. US, Springer, 2017.

[3] M. Chihaoui, A. Elkefi, W. Bellil, and C. B. Amar. "A Survey of 2D Face Recognition Techniques." Computers, vol. 5, no. 4, pp. 21, 2016.

[4] R. Jafri and H. R. Arabnia. "A survey of face recognition techniques." Jips, vol. 5, no. 2, pp. 41-68, 2009.

[5] T. Matthew, and A. Pentland. "Face recognition using eigenfaces." In Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991, pp. 586-591.

[6] T. Matthew, and A. Pentland. "Eigenfaces for recognition." Journal of cognitive neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[7] G. Shakhnarovich, and M. Baback, Face recognition in subspaces. In Handbook of Face Recognition, London, Springer, 2011, pp. 19-49.

[8] Y. Jian, D. Zhang, A. Frangi, and J. Yang. "Two-dimensional PCA: a new approach to appearance-based face representation and recognition." IEEE transactions on pattern analysis and machine intelligence, vol. 26, no. 1, pp. 131-137, 2004.

[9] M. Stewart, J. R. Movellan, and T. J. Sejnowski. "Face recognition by independent component analysis." IEEE Transactions on neural networks, vol. 13, no. 6, pp. 1450-1464, 2002.

[10] B. Peter, J. P. Hespanha, and D. Kriegman. "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection." IEEE Transactions on pattern analysis and machine intelligence, vol. 19, no. 7, pp. 711-720, 1997.

[11] H. Heiko. "Kernel PCA for novelty detection." Pattern Recognition, vol. 40, no. 3, pp. 863-874, 2007.

[12] L. Juwei, K. N. Plataniotis, and A. N. Venetsanopoulos. "Face recognition using kernel direct discriminant analysis algorithms." IEEE Transactions on Neural Networks, vol. 14, no. 1, pp. 117-126, 2003.

[13] B. R. Francis, and M. I. Jordan. "Kernel independent component analysis." Journal of machine learning research, vol. 3, no. 1, pp. 1-48, Jul. 2002.

[14] Y. Ming-Hsuan. "Face recognition using extended isomap." In Proc International Conference on Image Processing, 2002.

[15] L. Martin, J. C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Wurtz, and W. Konen. "Distortion invariant object recognition in the dynamic link architecture." IEEE Transactions on computers, vol. 42, no. 3, pp. 300-311, 1993.

[16] W. Laurenz, N. Krüger, N. Kuiger, and C. Von Der Malsburg. "Face recognition by elastic bunch graph matching." IEEE Transactions on pattern analysis and machine intelligence, vol. 19, no. 7, pp. 775-779, 1997.

[17] B. Roberto, and T. Poggio. "Face recognition: Features versus templates." IEEE transactions on pattern analysis and machine intelligence, vol. 15, no. 10, pp. 1042-1052, 1993.

[18] L. T. Sing. "Image representation using 2D Gabor wavelets." IEEE Transactions on pattern analysis and machine intelligence, vol. 18, no. 10, pp. 959-971, 1996.

[19] L. G. David. "Distinctive image features from scale-invariant keypoints." International journal of computer vision, vol. 60, no. 2, pp. 91-110, 2004.

[20] T. Ahonen, A. Hadid, and M. Pietikäinen. "Face recognition with local binary patterns." In Proc European conference on computer vision, 2004, pp. 469-481.

[21] N. V. Ara, and M. H. Hayes. "Face detection and recognition using hidden Markov models." In Proc International Conference on Image Processing, 1998, pp. 141-145.

[22] A. Kar, P. Bhattacharjee, M. Nasipuri, D. K. Basu, M. Kundu. “High performance human face recognition using gabor based pseudo hidden Markov model.” Int. J. Appl. Evol. Comput., 2013, pp. 81-102.

[23] J. S. Kais. "Face recognition system using PCA and DCT in HMM." Int. J. Adv. Res. Comput. Commun. Eng , vol. 4, pp. 13-18, 2015.

[24] M. Chihaoui, W. Bellil, A. Elkefi, and C. B. Amar. "Face recognition using HMM-LBP." Hybrid Intelligent Systems, pp. 249-258, 2016.

[25] H. Di, J. Sun, X. Yang, D. Weng, and Y. Wang. "3d face analysis: Advances and perspectives." In Proc Chinese Conference on Biometric Recognition, 2014, pp. 1-21.

[26] B. M. Alexander, M. M. Bronstein, R. Kimmel, and A. Spira. "3D face recognition without facial surface reconstruction." Technion-Computer Science Department Technical Report, 2003.

[27] B. Stefano, A. Del Bimbo, and P. Pala. "3D face recognition using isogeodesic stripes." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 12, pp. 2162-2177, 2010.

[28] B. W. Kevin, K. Chang, and P. Flynn. "A survey of approaches and challenges in 3D and multi-modal 3D+ 2D face recognition." Computer vision and image understanding, vol. 101, no. 1, pp. 1-15, 2006.

[29] H. Di, Y. Wang, and Y. Wang. "A robust method for near infrared face recognition based on extended local binary pattern." In Proc International Symposium on Visual Computing Conference, 2007, pp. 437-446.

[30] G. Friedrich, and Y. Yeshurun. "Seeing people in the dark: face recognition in infrared images." In Proc. International Workshop on Biologically Motivated Computer Vision, 2002, pp. 348-359.

[31] J. A. Laszlo, H. Hashimoto, and T. Kubota. "Robust Facial Expression Recognition Using Near Infrared Cameras." JACIII, vol. 16, no. 2, pp. 341-348, 2012.

[32] B. Stephen. "Deep learning and face recognition: the state of the art." In Proc. SPIE Defense+ Security, 2015.

[33] N. Jiquan, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng. "Multimodal deep learning." In Proc. The 28th international conference on machine learning (ICML-11), 2011, pp. 689-696.

[34] S. Yi, Y. Chen, X. Wang, and X. Tang. "Deep learning face representation by joint identification-verification." Advances in neural information processing systems, pp. 1988-1996, 2014.

[35] P. M. Omkar, A. Vedaldi, and A. Zisserman. "Deep Face Recognition." BMVC, vol. 1, no. 3, pp. 1-12, 2015.

[36] L. Yann, Y. Bengio, and G. Hinton. "Deep learning." Nature, vol. 521, no. 7553, pp. 436-444, 2015.

[37] T. Yaniv, M. Yang, M. Ranzato, and L. Wolf. "Deepface: Closing the gap to human-level performance in face verification." In Proc. The IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701-1708.

[38] S. Florian, D. Kalenichenko, and J. Philbin. "Facenet: A unified embedding for face recognition and clustering." In Proc. The IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 815-823.

[39] B. Amos, B. Ludwiczuk, and M. Satyanarayanan. “OpenFace: A general-purpose face recognition library with mobile applications.” Technical report, CMU-CS-16-118, CMU School of Computer Science, 2016.

[40] H. A. Khan, M. Al-Akaidi and R. W. Anwar. “Linear Discriminant Classifier Ensemble for Face Recognition.” In Proc. of MESM’ 2014.

[41] The AT & T Database of Faces. Internet: http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html [19 February 2017].

[42] The Yale Database. Internet: http://vision.ucsd.edu/content/yale-face-database [23 February 2017].

[43] N. Dalal and B. Triggs. “Histograms of oriented gradients for human detection.” In Proc IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), 2005, pp. 886-893.

[44] I. Daubechies. “Ten lectures on wavelets.” Society for industrial and applied mathematics, 1992.

[45] R. C. Gonzalez, and R. E. Woods. Digital Image Processing, Pearson International Edition, 2008.

[46] Z. Zhou, J. Wu, and W. Tang. “Ensembling neural networks: many could be better than all.” Artificial Intelligence, vol. 137, no. 1-2, pp. 239-263, 2002.