Feature Extraction and Classification Technique for Multi-Algorithm Facial Recognition System

Kehinde A. Sotonwa[1], Oluwashina A. Oyeniran[2]

[1]*Department of Computer Science and Information Technology, Bells University of Technology, Ota, Ogun State, Nigeria*
[2]*Department of Computer Science and Information Technology, Bowen University, Iwo, Osun State, Nigeria*

Abstract: **This study presents a multi-algorithm approach for Facial Recognition System (FRS) feature extraction and classification technique, so as to eradicate insubstantial accuracy. The system frontend was implemented using C# while Microsoft SQL Server was used for the backend; also, IBM SPSS was used for the statistical analysis of the data. Principal Component Analysis (PCA) with Histogram of Oriented Gradient (HOG) was combined for feature extraction while Artificial Neural Network (ANN) combined with Support Vector Machine (SVM) for classification. Five hundred registered and fifty unregistered facial images were evaluated on the standard biometric to verify the image. Statistically, there is no significant relationship between PCA and ANN, HOG and SVM for False Acceptance Rate (FAR) while there is significant relationship between PCA and ANN, HOG and SVM for False Rejected Rate (FRR). Moreover, correlations between PCA with ANN and HOG with SVM using Chi-Square and t-test indicate that PCA with ANN show a more reliable measure than HOG with SVM.**

***Keywords*: FRS, Feature Extraction, C#, Microsoft SQL Server, Multi-Algorithm.**

I. INTRODUCTION

It is expected today that an individual who wants to validate himself for a service must have a token and/or password, for example, identity card, ATM card, driving license, international passport and so on are continuously requested for and carrying different cards and remembering passwords for different services is a momentous concern for individuals and organizations [1]. A secure and effective identity management system plays an imperative function in the successful deployment of an attendance management system [2]. Biometrics technologies verify identity through characteristics such as fingerprints, faces, irises, retinal patterns, palm prints, voice, hand-written signatures, and so on [3].

These techniques, which use physical data, are receiving attention as a personal authentication method that is more convenient than conventional methods such as a password or ID cards because it uses data taken from measurements and such data is unique to the individual and remains so throughout one's lifetime [4]. Human beings had recognition capabilities that are unparalleled in the modern computing era and have always had the ability to recognize and distinguish between face features and with advent of machine and deep learning.

The computers have been shown to have same ability to recognize and distinguish between face features. In the mid of 1960s, scientists began work on using the computer to recognize human faces, since that time, the facial recognition software has come a long way that the government and private corporations start to use the facial recognition system. In recent years, facial recognition has received substantial attention from researchers in biometrics, pattern recognition, and computer vision communities [5]. The machine learning and computer graphics communities are also increasingly involved in facial recognition. This common interest among researchers working in diverse fields is motivated by our remarkable ability to recognize people and the fact that human activity is a primary concern both in everyday life and in cyberspace.

Besides, there are large numbers of commercial, security, and forensic applications requiring the use of facial recognition technologies [5]. Face recognition is one of the most relevant applications of image analysis. It is a true challenge to build an automated system with equal human ability to recognize faces. Although humans are quite good identifying known faces but are not very skilled when dealing with a large number of unknown faces [5]. The computers, with an almost limitless memory and computational speed, should overcome these human limitations. Thus, face recognition remain a demanded technology as there are many different industrial areas that are interested in what it could offer.

Some examples include video surveillance, human-machine interaction, photo cameras, virtual reality or law enforcement [6]. This multidisciplinary interest instigates research and attracts interest from diverse disciplines. Therefore, it is not a problem that is restricted to computer vision research. Face recognition is a relevant subject in pattern recognition, neural networks, computer graphics, image processing and psychology [7]. In the global community, various facial recognition frameworks have being proposed and successfully implemented, many of which are of one limitation or the other. Limitations like; the issue of high false acceptance and rejection rate, low recognition accuracy, illumination, posing and expression among others. Hence, this study proposed a multi-algorithm approach for facial

recognition system feature extraction and classification technique, so as to eradicate insubstantial accuracy.

II. PROBLEM STATEMENT

Although human are quite good identifying known faces but are not yet skilled when dealing with a large number of unknown faces [5]. The computers, with an almost limitless memory and computational speed should overcome these human limitations. Thus face recognition remain a demanded technology as there are many different industrial areas that are interested in what it could offer. Some examples include video surveillance, human machine interaction, photo cameras, virtual reality or law enforcement [6]. This multidisciplinary interest instigates research and attracts interest from diverse discipline [7]. In the global community, various facial recognition frameworks have being proposed and successfully implemented, many of which are with one limitation or the other, for example the issue of high false acceptance and rejection rate, low recognition accuracy, illumination, posing and expression among others,. Hence, this study proposed a multi algorithm approach for facial recognition system feature extraction and classification technique so as to eradicate insubstantial accuracy.

III. RELATED WORK

In order to solve the problem of low accuracy of face recognition under non – restrictive conditions, a new method of face recognition based on Haar feature classifier, HOG feature extraction and fast-PCA dimension reduction was proposed by [8] in their study titled Face Recognition Based on HOG and Fast PCA Algorithm. Firstly, the Haar feature classifier is used to extract the background interference data at the same time in the original data preprocessing stage. Then, the feature data of the face is extracted by the method of HOG feature extraction. Then, the extracted data PCA algorithm is reduced the size of the final use for the training and testing of the amount of data. Finally, the use of SVM algorithm is to identify and identify the face. It is verify the effectiveness of the method with the experimental results.

Reference [9] also implemented a facial recognition system using a global-approach to feature extraction based on Histogram-Oriented Gradient. They extracted the feature vectors for various faces from the AT&T and Yale databases and used them to train a binary-tree structure SVM learning model. Running the model on both databases resulted in over 90% accuracy in matching the input face to the correct person from the gallery.

They noted one of the shortcomings of using a global approach to feature extraction, which is that a model trained using a feature vector of the entire face instead of its geometrical components makes it less robust to angle and orientation changes. However, when the variation in facial orientation is not large, the global-approach is still very accurate and simpler to implement than component-based approaches.

Moreover, [10] proposed system consists of a camera that captures the images of the classroom and sends it to the image enhancement module. After enhancement the image comes in the Face Detection and Recognition modules and then the attendance is marked on the database server. At the time of enrolment templates of face images of individual students are stored in the Face database. Here all the faces are detected from the input image and the algorithm compares them one by one with the face database.

If any face is recognized the attendance is marked on the server from where anyone can access and use it for different purposes. This system uses a protocol for attendance. A time table module is also attached with the system which automatically gets the subject, class, date and time. Teachers come in the class and just press a button to start the attendance process and the system automatically gets the attendance without even the intensions of students and teacher.

In this way a lot of time is saved and this is highly secure process; no one can mark the attendance of other. Attendance is maintained on the server so anyone can access it for it purposes like administration, parents and students themselves. Camera takes the images continuously to detect and recognize all the students in the classroom. In order to avoid the false detection we are using the skin classification technique. Using this technique enhances the efficiency and accuracy of the detection process but the system has two major flaws, which are:

(i) It can only detect face from a limited distance.
(ii) The system don't recognized properly in poor light so may give false results.

This research work proposed to use Principal Component Analysis combined with Histogram of Oriented Gradient for the feature extraction and Artificial Neural Network combined with Support Vector Machine for the classification and the recognition accuracy of the system will be determined by employing it on some public parameters like Execution Time (ET), False Acceptance Rate (FAR), False Rejection Rate (FRR), and Total Percentage Accuracy (TPA) (see chapter four for formulae).

IV. METHODOLOGY

Humans are *too good* at recognizing faces and end up seeing faces in everyday objects. Computers are not capable of this kind of high-level generalization (*at least not yet*), so we have to teach them how to do each step in this process separately. To do this, we need to build a pipeline where we solve each step of face recognition separately and pass the result of the current step to the next step. In other words, we will chain together several machine learning algorithms: find faces in an image, analyze facial features, compare against known faces and make a prediction.

In order to accomplish the objectives of this study, a model was formulated for the design (see figure 1 below). The developed model was implemented using C#. Principal Component Analysis (PCA) combined with Histogram of Oriented Gradient (HOG) was used for feature extraction and Artificial Neural Networks (ANN) combined with Support Vector Machine (SVM) was used for classification.

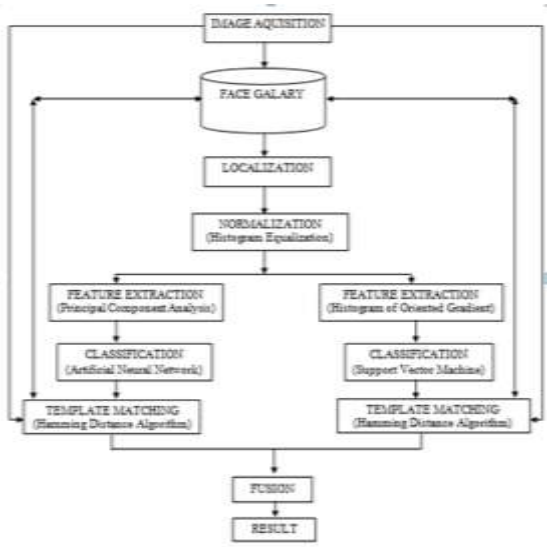

Figure 1: The Proposed Framework for Computerized Student Attendance Management System

V. RESULT AND DISCUSSION

In order to evaluate the performance of the system, intensive and extensive testing was conducted, where five hundred and fifty Facial images were collected and subjected to test (five hundred of which are used for enrollment and the remaining fifty are not registered). Thus, this study employed the metrics below for the evaluation of the system.

Table 1: Comparison of the Algorithms

Threshold values	PCA with ANN		HOD with SVM	
	FAR	FRR	FAR	FRR
1.2	0.3	3.5	0.2	3.1
1.4	0.4	0.5	0.4	0.2
1.6	0.2	0.3	0.2	0.5
1.8	1.9	0.9	2.2	0.6

A. Hypothesis (I):

H_0: There is no significant relationship between Principal Component Analysis with Artificial Neural Network (FAR) and Histogram of Oriented Gradient with Support Vector Machine (FAR)

Decision rule: Reject H_0 if p-value is less than or equal to significant value

Table 2: Correlations between PCA with ANN and HOG with SVM (FAR)

		PCA with ANN (FAR)	HOG with SVM (FAR)
PCA with ANN (FAR)	Pearson Correlation	1	0.732
	Sig. (2-tailed)		0.159
	N	5	5
HOG with SVM (FAR)	Pearson Correlation	0.732	1
	Sig. (2-tailed)	0.159	
	N	5	5

The findings of this study with respect to Table 2 indicate the correlations between Principal Component Analysis with Artificial Neural Network (FAR) and Histogram of Oriented Gradient with Support Vector Machine (FAR). Since p-value (0.159) is greater than the significant value (0.005), it is therefore concluded that there is no significant relationship between Principal Component Analysis with Artificial Neural Network (FAR) and Histogram of Oriented Gradient with Support Vector Machine (FAR)

B. Hypothesis (II):

H_0: There is no significant relationship between Principal Component Analysis with Artificial Neural Network (FRR) and Histogram of Oriented Gradient with Support Vector Machine (FRR)

Decision rule: Reject H_0 if p-value is less than or equal to significant value

Table 3: Correlations between PCA with ANN and HOG with SVM (FRR)

		PCA with ANN (FAR)	HOG with SVM (FAR)
PCA with ANN (FAR)	Pearson Correlation	1	0.970
	Sig. (2-tailed)		0.006
	N	5	5
HOG with SVM (FAR)	Pearson Correlation	0.970	1
	Sig. (2-tailed)	0.006	
	N	5	5

The findings of this study with respect to Table 3 indicate the correlations between Principal Component Analysis with Artificial Neural Network (FRR) and Histogram of Oriented Gradient with Support Vector Machine (FRR). Since p-value (0.006) is less than the significant value (0.005), it is therefore concluded that there is significant relationship between Principal Component Analysis with Artificial Neural Network (FRR) and Histogram of Oriented Gradient with Support Vector Machine (FRR)

C. Hypothesis (III):

H_0: Principal Component Analysis with Artificial Neural Network is not reliable than Histogram of Oriented Gradient with Support Vector Machine

Decision rule: Reject H_0 if calculated value is greater than tabulated value (using chi-square)

Table 4: Correlations between PCA with ANN and HOG with SVM using Chi-Square

	PCA WIH ANN	HOG WIH SVM
Chi-Square	1.200[a]	2.800[a]
df	7	7
Asymp. Sig.	.991	.903

The findings of this study with respect to Table 4 indicate the correlations between Principal Component Analysis with Artificial Neural Network and Histogram of Oriented Gradient with Support Vector Machine (Using Chi-Square). Since the calculated value (0.99) and (0.903) are less than the tabulated value (3.247), then we do not reject H_0. It is therefore concluded that PCA with ANN is more reliable than HOG with SVM.

D. Hypothesis (IV):

H_0: Principal Component Analysis with Artificial Neural Network is not reliable than Histogram of Oriented Gradient with Support Vector Machine.

Decision rule: Reject H_0 if calculated value of Histogram of Oriented Gradient with Support Vector Machine is greater than calculated value of Principal Component Analysis with Artificial Neural Network (using *t*-Test)

Table 5: Mean Comparison between PCA with ANN and HOG with SVM Using *t*-Test

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
				0.9700	Lower	Upper
PCA WITH ANN	2.948	9	0.016		0.226	1.714
HOG WITH SVM	2.557	9	0.031	0.8200	0.095	1.545

Table 5 also provides the results of the *t*-test, which tested the differences between the mean of Principal Component Analysis with Artificial Neural Network and Histogram of Oriented Gradient with Support Vector Machine. Since the calculated value (0.016) of Principal Component Analysis with Artificial Neural Network is less than the calculated value (0.031) of Histogram of Oriented Gradient with Support Vector Machine, then we do not reject H_0. It is therefore concluded that PCA with ANN is more reliable than HOG with SVM.

(1) *Total Execution Time:* The recognition time for the input image is approximately 2.00 seconds while it takes approximately 4.00 seconds to verify the image.

(2) *False Acceptance Rate (FAR):* The false acceptance rate, or FAR, is the measure of the likelihood that the biometric security system will incorrectly accept an access attempt by an unauthorized user. A system's FAR typically is stated as the ratio of the number of false acceptances divided by the number of identification attempts. For the purpose of this study, five hundred facial images were registered and verified, in which none of the unregistered image were accepted. Hence;

$$FAR = \frac{Number\ of\ Acceptance\ Image}{Number\ of\ Registered\ Image} * 100 \qquad (1)$$

*FAR = 1/100*100 = .2%*

(3) *False Rejection Rate (FRR):* The false recognition rate, or FRR, is the measure of the likelihood that the biometric security system will incorrectly reject an access attempt by an authorized user. A system's FRR typically is stated as the ratio of the number of false recognitions divided by the number of identification attempts. For the purpose of this study, one hundred facial images were registered and verified, in which two of the registered image were rejected. Hence;

$$FRR = \frac{Number\ of\ False\ Rejected\ mage}{Number\ of\ Registered\ Imag} * 100 \qquad (2)$$

*FRR = 15/500*100 = 3%*

(4) *Recognition Accuracy:* Recognition accuracy of a biometric system also refers to as average correct recognition rate, which is the total percentage of the correct recognition of the system. Recognition accuracy is defined as follows:

Recognition accuracy = (100 – (FAR + FRR)) % (3)

Recall: FAR = 0.2% while FRR = 3%

Recognition accuracy = (100 – (0.2 + 3)) %

Recognition accuracy = ((100 – 3.2)) %

Recognition accuracy = 96.8%

(5) *Equal Error Rate (EER):* Equal error rate (EER) is a biometric security system algorithm metric used to predetermine the threshold values for its false acceptance rate and its false rejection rate. When the rates are equal, the common value is referred to as the *equal error rate*. The value indicates that the proportion of false acceptances is equal to the proportion of false rejections. The lower the equal error rate value, the higher the accuracy of the biometric system. Equal error rate may also be referred to as a Crossover Rate or Crossover Error Rate (CER). The equal error rate or *this* proposed system is calculated to be 0.3% (see figure 2 below).

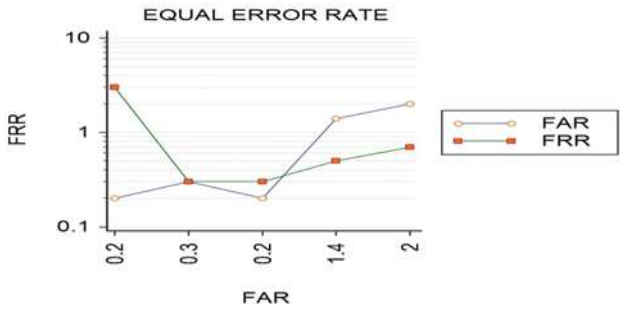

Figure 2: Equal Error Rate

VI. CONCLUSION

The results of this study indicate that multiple feature extraction algorithm and classifier improves the performance of a biometric authentication with better accuracy and reduce error rate. Moreover, the statistical analysis reveals that Principal Component Analysis combined with Artificial Neural Network reduces False Rejection Rate (FRR), also, the result of the statistical analysis indicate that PCA with ANN is more reliable than HOG with SVM. Hence, this study recommends that future researchers should venture into ensemble of algorithms to take care major drawbacks in face recognition system

REFERENCES

[1]. Kadry S. and Smaili M. (2010). Wireless Attendance Management System based on Iris Recognition. *Scientific Research and Essays Vol. 5*(12), pp. 1428-1435

[2]. Khan B., Khan M. K. and Alghathbar K. S. (2010).Biometrics and Identity Management for Homeland Security Applications in Saudi Arabia.*African Journal of Business Management* Vol. 4(15), pp. 3296-3306.

[3]. Akinduyite, C. O, Adetunmbi, A. O, Olabode, O. O, Ibidunmoye, E. O (2013) Fingerprint-Based Attendance Management System. *Journal of Computer Sciences and Applications, Vol. 1*, No. 5, Pp. 100-105

[4]. Shoewu, O., Olaniyi, O .M. and Lawson, A. (2011).Embedded Computer-Based Lecture Attendance Management System.*African Journal of Computing and ICT*.Vol 4, No. 3. P 27- 36

[5]. Asmahan, M. A. (2014). 'Face Recognition Techniques - An Evaluation Study' *International Journal of Advanced Networking and Applications Volume: 6* Issue: 4 Pages: 2393-2397 (2015) ISSN: 0975-0290

[6]. Shirodkar, M., Sinha, V., Jain, U. and Nemade, B. (2015).Automated Attendance Management System using Face Recognition.*International Journal of Computer Applications (0975 – 8887) International Conference and Workshop on Emerging Trends in Technology*

[7]. Chang, H., Koschan, A., Abidi, B. and Abidi, M. (2010). Fusing Continuous Spectral Images for Face Recognition under Indoor and Outdoor Illuminants. *Machine Vision and Applications*, 21(2):201–215, February 2010.

[8]. Li, S. and Lin, L. (2017). Face Recognition Based on HOG and Fast PCA Algorithm *Proceedings of the fourth Euro-China Conference on Intelligent Data Analysis and Applications* http://www.springer.com/cda/content/document/cda_downloaddocument/9783319685267-c2.pdf?SGWID=0-0-45-1619846-p181166638 retrieved on November 23rd, 2017

[9]. Arun, G. and Peter, C. (2017). Face Recognition using Machine Learning http://noiselab.ucsd.edu/ECE285/FinalProjects/Group7.pdf*(retrieved on November 23rd, 2017)*

[10]. Navesh, P., Shubham, T., Vaibhav, R., Vishal, Y., Parag, D. and Gaurav, F. (2017).Automatic Attendance System by Using Face Recognition. *International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 04 Issue: 04*https://www.irjet.net/archives/V4/i4/IRJET-V4I4552.pdf*(retrieved on November 23rd, 2017)*