International Journal of Engineering and Technology Volume 2 No. 11, November, 2012

Implementation of an Improved Facial Recognition Algorithm in a Web based Learning System

Adeolu Olabode Afolabi[1] , Rotimi Adagunodo[2]
[1]Department of Computer Science and Engineering,Ladoke Akintola University of Technology, Ogbomoso. Nigeria
[2]Department of Computer Science and Engineering, Obafemi Awolowo University Ile Ife, Nigeria

ABSTRACT

The study focuses on proffering solution to some identified data insecurity problems in software development using Web-based learning system as a test bed, by development of an hybrid crypto-biometric security system, and the use of an enhanced eigen-based facial recognition algorithm.

The methodology is by implementation of an optimized principal component analysis eigen facial recognition algorithm for black faces using matlab. A comparative analysis of performance of the optimized principal component analysis (OPCA) and (PCA) principal component analysis is done and it was found out that OPCA performed better than PCA. Also a web based learning system using Hypertext Pre-processor (PHP), Scripting Language for the Web-based pages, Asynchronous JavaScript and XML (AJAX) is developed as a test bed for the crypto biometric system.

With this work a prototype for the secured Web-based learning infrastructure is developed and its contextual framework, also an optimized principal component analysis algorithm for black face recognition evolve as contributions to knowledge. hence it will foster indigenization of electronic learning technology which will adequately address the related challenges in the phenomenon of system security in terms of confidentiality and integrity of the system.

Keywords: *Web based Learning, Principal Component Analysis, Face Recognition*

1. INTRODUCTION

Electronic learning (Web-based learning) is the online delivery of information through computer networks for the purpose of education, training, knowledge management or performance management. Also, it involves the use of Internet Technology, as a platform for making education available to remote communities. This may not be feasible using conventional teaching methods within a geographical location [Oweston, 2005]. The application of Technology in Education has been adopted in developed countries of the world, and this has improved the quality of life of the beneficiaries as well as reduced the cost of learning [Tumi, 2005].

Several Web-based learning portals have been developed in recent times for the purpose of facilitating their operations, however the portals have been found to have some inadequacies and limitations in terms of functionality and performance. These are largely due to some attendant problems such as insecurity of the system (lack of authentication of the genuine users, loss of data content, performance evaluation of the students and proper deployment of needed courseware) as well as the weakness of the system which often times may not be sufficiently robust to withstand the attendant problems.

Arising from the identified problems of the new technology, this work intends to demonstrate the effectiveness of a unique technique that will address insecurity issue that seems to hamper effective operation of various Web-based learning systems. There is need for a secured E- Learning portal that would guarantee learning within the virtual classroom.

2. METHODOLOGY

The methodology involves the development of crypto-biometric hybrid system by implementing an optimized (optimized principal component analysis) eigen facial recognition algorithm for black faces using matlab. As a test bed for this security paradigms, a web based learning system using Hypertext Pre-processor (PHP), Scripting Language for the Web-based pages, Asynchronous JavaScript and XML (AJAX) to enhance Chatting and Macromedia Flash and other relevant application like Standard Query Language Database Management System for storing data will be developed.

The perceived system is a feedback system that can handle the performance evaluation of student and assessment online. A web cast of the developed Web-based learning portal will facilitates online interaction between the learners and the instructors, the user will gain access via authentication by facial recognition features and by the use of appropriate encryption key to the data within the framework of Web-based learning environment. The various techniques are used so that the courseware could be assessed and sent to the learner in a virtual classroom for learners' feedback in a fulfilled security conditions.

3. FACE RECOGNITION

Face recognition is a biometric technology that uses an image or series of images either from a camera or photography for recognition of a person. Unlike other biometric technologies, face recognition is a passive biometric and does not require a person's cooperation [Pentland,1997]. It can recognize people from a distance without them realizing they are being analyzed; face recognition is completely different in appearance as a result of race or gender difference and is a highly robust biometrics.

This biometrics technology has been used extensively throughout the world over the last three to five years in industries like Banking, Gaming, Healthcare, Law enforcement, Customs and Excise. The technology has proven extremely successful and can be regarded as the fastest growing biometrics in the world.

The technology use an algorithm called LFA (Local Feature Analysis) to identify and derive a representation in terms of the spatial relationship between irreducible local features, or 'nodal point' on the face.

4. FACE RECOGNITION ALGORITHMS

In this work facial recognition is going to be used as the biometric feature in the Web-based learning system that is to be developed based on optimized (PCA) principal component analysis algorithm.

4.1 Principal Component Analysis (PCA)

[Sirovich & Kirby1987] first used PCA to efficiently represent pictures of human faces. [Turk & Pentland 1991] presented the well-known Eigenfaces method for face recognition in 1991. Since then, PCA has been widely investigated and has become one of the most successful approaches in face recognition.

In all previous PCA-based face recognition technique, the 2D face image matrices must be previously transformed into 1D image vectors. The resulting image vectors of faces usually lead to a high dimensional image vector space, where it is difficult to evaluate the covariance matrix accurately due to its large size and the relatively small number of training samples.

With this background there is a need for improved PCA that will evaluate the covariance matrix accurately and more easily also that will require less time to determine the corresponding eigenvectors. This is further explicitly established in this work in order to demonstrate the effectiveness of improved methods.

4.2 Eigenface - Based Facial Recognition

The task of facial recognition is discriminating input signals (image data) into several classes (persons). The input signals are highly noisy (e.g. the noise is caused by differing lighting conditions, pose and other factors), yet the input images are not completely random and in spite of their differences there are patterns which occur in any input signal. Such patterns, which can be observed in all signals, could be in the domain of facial recognition - the presence of some objects (eyes, nose, mouth) in any face as well as relative distances between these objects. These characteristic features are called *eigenfaces* in the facial recognition domain (or *principal components* generally). They can be extracted out of original image data by means of a mathematical tool called *Principal Component Analysis.*

By means of PCA one can transform each original image of the training set into a corresponding eigenface. An important feature of PCA is that one can reconstruct any original image from the training set by combining the eigenfaces. Therefore one could say that the original face image can be reconstructed from eigenfaces if one adds up all the eigenfaces (features) in the right proportion. Each eigenface represents only certain features of the face, which may or may not be present in the original image. If the feature is present in the original image to a higher degree, the share of the corresponding eigenface in the "sum" of the eigenfaces should be greater. On the contrary, if the particular feature is not (or almost not) present in the original image, then the corresponding eigenface should contribute a smaller (or not at all) part to the sum of eigenfaces. So, in order to reconstruct the original image from the eigenfaces, one has to build a kind of weighted sum of all eigenfaces. That is, the reconstructed original image is equal to a sum of all eigenfaces, with each eigenface having a certain weight. This weight specifies, to what degree the specific feature (eigenface) is present in the original image.

If one uses all the eigenfaces extracted from original images, one can reconstruct the original images from the eigenfaces *exactly*. But one can also use only a part of the eigenfaces. Then the reconstructed image is an approximation of the original image. However, one can ensure that losses due to omitting some of the eigenfaces can be minimized. This happens by choosing only the most important features (eigenfaces). Omission of eigenfaces is necessary due to scarcity of computational resources.

4.3 Implementation

The test bed for this project is a web based application that makes use of web clients, business components, and application server as well as a web server. It is written entirely in Java and is based on the J2 Enterprise Edition platform. The software consists of two major components, the web application and the Face recognitions module. The general overview of the software is shown in Figure 1.

5. SECURED WEB-BASED LEARNING ARCHITECTURAL FRAMEWORK

The architectural framework of the secured Web-based learning is shown in figure 3.4 it describes the architectural framework graphically for the system being developed, the architectural design shows the various components of the development.

6. FACE RECOGNITION AND VERIFICATION MODULE

The face recognition module consists of two components: Face capture module and the face verification module. Face Capture Module: This module is responsible for the following:

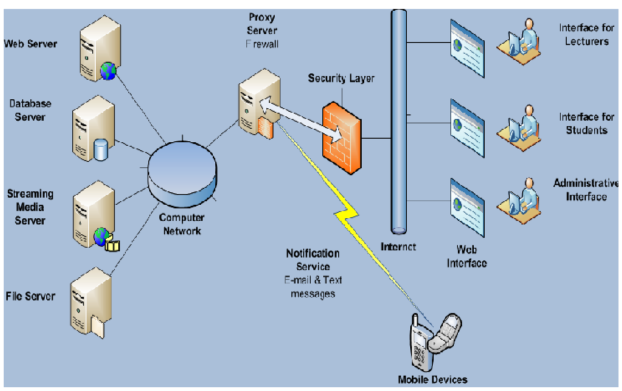

Figure 1 Secured Web-based learning architectural Framework

7. IMPLEMENTATION OF EIGENFACES WITH PCA

In this section, the original scheme for determination of the eigenfaces using PCA will be presented. The algorithm described below is the basis of the one being optimized within the scope of this work as a variation of the one outlined here.

Step 1: Prepare the data

In this step, the faces constituting the training set (Γ_i) should be prepared for processing.

Step 2: Subtract the mean

The average matrix has to be calculated, then subtracted from the original faces (Γ_i) and the result stored in the variable Φ_i:

$$\Psi = \frac{1}{M}\sum_{n-1}^{M}\Gamma_n \quad (1)$$

$$\Phi_i = \Gamma_i - \Psi \quad (2)$$

Step 3: Calculate the covariance matrix

In the next step the covariance matrix *C* is calculated according to

$$C = \frac{1}{M}\sum_{n-1}^{M}\Phi_n\Phi_n^T \quad (3)$$

Step 4: Calculate the eigenvectors and eigenvalues of the covariance matrix

In this step, the eigenvectors (eigenfaces) u_i and the corresponding eigenvalues λ_i should be calculated. The eigenvectors (eigenfaces) must be normalised so that they are unit vectors, i.e. of length 1.

Step 5: Select the principal components

From *M* eigenvectors (eigenfaces) u_i, only *M'* should be chosen, which have the highest eigenvalues. The higher the eigenvalue, the more characteristic features of a face does the particular eigenvector describe. Eigenfaces with low eigenvalues can be omitted, as they explain only a small part of characteristic features of the faces. After *M'* eigenfaces u_i are determined, the "training" phase of the algorithm is finished.

7.1 Distinguished features of PCA (Principal Components Analysis)

Principal Components Analysis works with its consideration for the nodal points in the face region through dimensionality reduction, these points are

(i) distance between eyes
(ii) the shape of the cheekbones
(iii) depth of the eye sockets
(iv) width of the nose.

These are distinguishable features of every human and peculiar to each man's face, these are to be considered in the algorithm to be improved upon for recognition of black faces. The major optimization of the algorithm is going to be carried out on the illumination and extraction of face features stages of the algorithm in order for better acceptance rate of recognition of black faces.

7.2 Improvement of the Original Algorithm

Roweis S (2002) posited that for learning the principal components of a dataset the algorithm does not require computing the sample covariance and has a complexity limited by operations where the number of leading eigenvectors is to be learned.

Another shortcoming of standard approaches to PCA is that it is not obvious how to deal properly with missing data. Most of the methods of PCA cannot accommodate missing values and so incomplete points must either be discarded or completed using a variety of ad-hoc interpolation methods.

Finally, the PCA model itself suffers from a critical flaw which is independent of the technique used to compute its parameters: it does not define a proper probability model in the space of inputs, this is because the density is not normalized within the principal subspace.

With the algorithm described in section above, the covariance matrix C in step 3 (see equation 3.10) has a dimensionality of $N^2 \times N^2$, so one would have N^2 eigenfaces and eigenvalues. For a 256 × 256 image that means that one must compute a 65, 536 × 65, 536 matrix and calculate 65,536 eigenfaces. Computationally, this is not very efficient as most of those eigenfaces are not useful for the task.

So, the step 3 and 4 is replaced by the scheme proposed by Turk and Pentland (1991a): where L is a $M \times M$ matrix, v are M eigenvectors of L and u are eigenfaces. Covariance matrix C is calculated using the formula $C = AA^T$, the original (inefficient) formula is given only for the sake of explanation of A. The advantage of this method is that one has to evaluate only M numbers and not N^2. Usually, $M \ll N^2$ as only a few principal components (eigenfaces) will be relevant. The amount of calculations to be performed is reduced from the number of pixels $N^2 \times N^2$ to the number of images in the training set (M).

In the step 5, the associated eigenvalues allow one to rank the eigenfaces according to their usefulness. Usually, only a subset of M eigenfaces, the M' eigenfaces with the largest eigenvalues is used.

$$C = \frac{1}{M}\sum_{n=1}^{M}\Phi_n\Phi_n^T = AA^T \qquad (4)$$

$$L = A^T A \quad L_{n,m} = \Phi_m^T\Phi_n \qquad (5)$$

$$u_l = \sum_{k=1}^{M} v_{lk}\Phi_k \quad l = 1,\ldots,M \qquad (6)$$

8. MATHEMATICAL ANALYSIS OF OPTIMIZED PRINCIPAL COMPONENT ANALYSIS (OPCA)

There is the need to address the issue of discrimination between the clusters in face recognition where the training data are labeled, ordinary PCA fails to accomplish this, as it is considered to be unsupervised techniques. The directions that maximize the scatter of the data might not be as adequate to discriminate between clusters. Therefore, new optimized PCA-based schemes which can correctly take into consideration data labeling, and makes the performance of recognition system better is developed. Experimental results show that the method achieves better performance in comparison with the traditional PCA method. In this section, a review on the basic notions, essential mathematical background and algorithms of PCA approaches that are needed for subsequent derivations of the optimized version are discussed. Images were converted into grey scale with pixel values between 0 (black) and 255 (white) because the algorithm requires two dimensional arrays for analysis, the recognition rates was tested with different number of training samples. kk (= 2,3,4,5) images of each subject are randomly selected from the database for training and the remaining images of each subject for testing. For each value of k , 30 runs are performed with different random partition between training set and testing set. And for each k training sample experiment, we tested the recognition rates with different number of dimensions , d ,time taken to train a face and time taken before an image could be identified. Different number of dimensions of pixel used are 45x45,60x60,75x75, 90x90 and 100x 100 fig 3.10a and 3.12 show the samples of these different dimensions of pixels used on blackfaces.

100x100 pixels

90x90 pixels

60x60 pixels

75x75 pixels

45x45 pixels

Fig 2. Different dimensions of pixels (cropped images) of black face

Table 1. Training of images using Optimised (OPCA) and PCA

Dimension of Pixels	Total number of images used	Total number of Unidentified Images PCA	Total number of Unidentified Images OPCA	% face recognition PCA	% face recognition OPCA (%)	Time to train face database PCA (sec)	Time to train face database OPCA (sec)	Time to identify known and unknown (sec) PCA	Time to known a unknowr (sec) OPCA
45x45	76	15	10	80.00	86.84	23.78	21.32	0.301	0.295
60x60	76	11	8	85.33	89.47	35.42	26.45	0.411	0.310
75x75	76	7	5	90.66	93.33	46.37	32.60	0.581	0.450
90x90	76	5	3	93.33	96.60	51.25	41.27	0.601	0.521
100x100	76	3	1	96.60	98.68	56.33	44.20	0.690	0.551

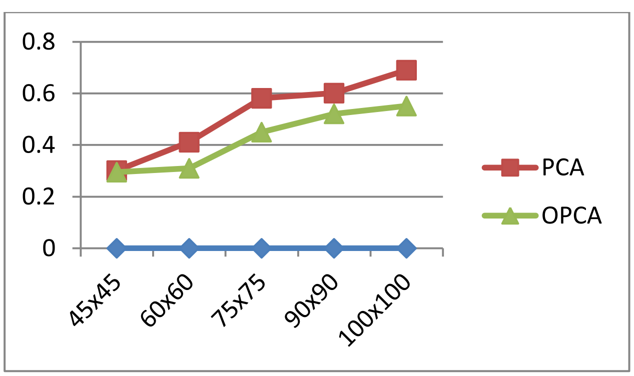

Fig 3 Time of Identification versus dimension of cropped images using PCA and OPCA

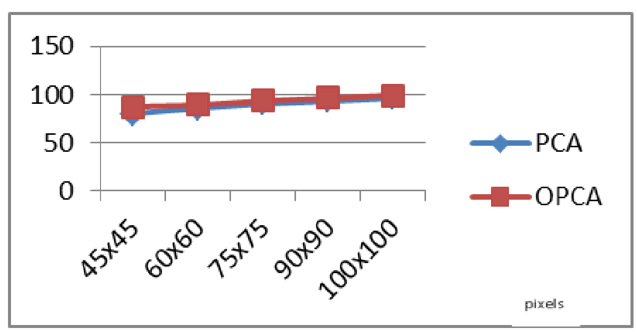

Fig 4 Percentage face recognition rate of PCA and OPCA

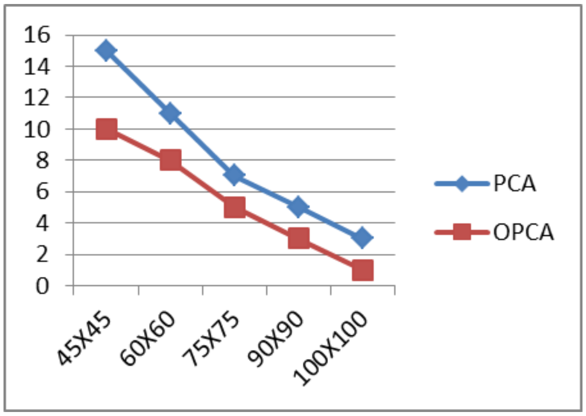

Fig 5. Comparison of unidentified images using PCA and OPCA algorithms

9. DISCUSSION OF RESULTS

The matlab code of both optimized principal component algorithm(OPCA) and PCA were tested on the Intel (R) Core TM Duo CPU of processor 1.83GHz and RAM of 3062MB Computer System. The mode of training was based on three main aspects of face recognition image acquisition, standardisation and dimensionality reduction. The training and testing were done on the cropped images stored in the database for recognition and identification as discussed previously.

Table 3.1 shows the result of the training done using PCA and OPCA algorithms with different dimensions of pixels in order to analyze and compare the performance of the two algorithms. Fig 3 shows the graphical representation of time of identification versus dimension of cropped images, it was discovered that the time taken for identification based on OPCA is faster than that of PCA. Also, Fig 4 shows the face recognition rate of PCA and OPCA and it was observed that the rate of OPCA is faster than that of PCA based on the data shown in Table 1 and this suggest that the optimized algorithm is better than the existing PCA. Fig 5 expressed the time taken to train faces on the database using different pixels of seventy six images of different individuals ,the results show that time taken to train faces with OPCA algorithms is faster than that of PCA. In all, the performance of OPCA is better considering the face recognition rate, time taken to train the face and time of identification.

9.1 Design and Discussion of Web Based Learning System

The software is to be written using JAVA programming language and MYSQL as the database here below are the various forms showing different functions.

9.1.1 Face Capture Module

The figure 6 shows the applet used to capture the face for storage and verification or for verification if the users face has already been stored.

After the successfully logs on, the face capture applet loads and the user is prompted to capture his face using the web camera as feed. On clicking the capture button, a preview of the shot taken is shown and the user can afford to either accept it or reject it to take another one.

The figure 7 shows the system when tested with three registered users' images and an outsider i.e. an unregistered individual. The pictures of the already identified users are shown in the figure. To complete the registration each user took three to four shots which formed the basis for comparison against the individual used as the test

Figure 6 shows the results obtained from capturing and submitting one of the registered faces. On capturing the face, a corresponding jpeg file is created and sent to the server for processing after which the result of the verification is sent to the applet. If an unregistered / unknown face is submitted the applet reports this as shown in figure 8 that is match not found for unknown user.

9.1.2 Administrator Module

In module the administrator has the privileges of adding users, adding or creating courses as well as assigning the courses to respective users. The site administrator has the privileges of creating users and courses as shown in figure 9. In addition, he is also responsible for assigning courses to user's designated as lecturers, each user must have a unique combination of first name and password. Also in fig 10, the administrator can also specify the course title and course code for the particular course to be created and creates as many as he needs to.

9.1.3 Lecturer Module

As a lecturer, a user can create and upload course materials as shown in the figure 11 In addition a lecturer can also carry out assessment in the form of a quiz, to do this a quiz is created after which the questions for the quiz are set. Each question is in multiple choice format. At the end of each question, the lecturer specifies the correct option and the questions are then saved, as shown in figure 12.

9.1.4 Student Module

If the student is not an authentic registered student the course material will appear in encrypted form before the student will now decrypt the material accordingly as shown in Figure 13. Here the student is allowed to view course material displayed by topic for each course registered for. Also, the student can take quizzes and as well register himself for other courses.

9.2 View Course Material

When a student logs into the system, he is shown a complete set of courses that he has registered for. Each course has a list of topics that have been created by the lecturers in charge. Each topic can be accessed through the link provided. Figure3.16 shows the encrypted view of the course material to the student.

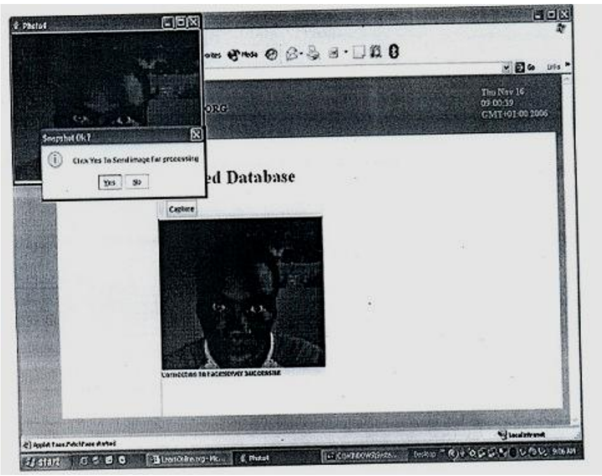

Fig 8 Face Capture Applet

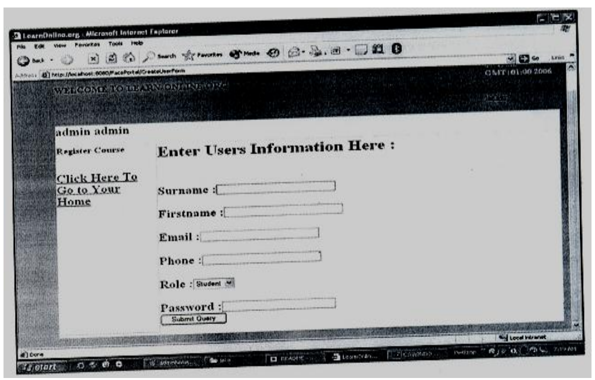

Fig 9 Create User

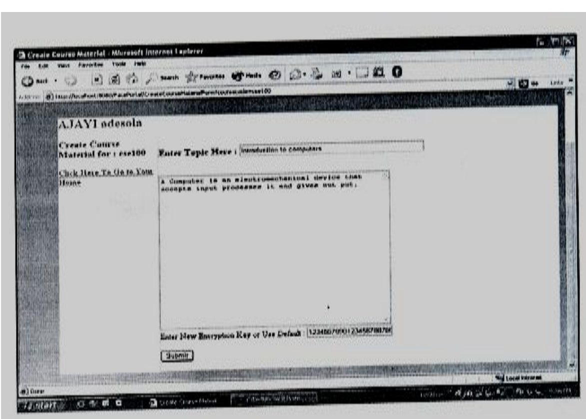

Fig 10 Creating Course Material

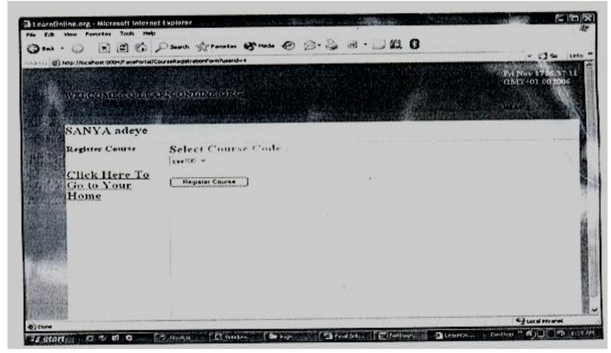

Fig. 11 Course Registration Window

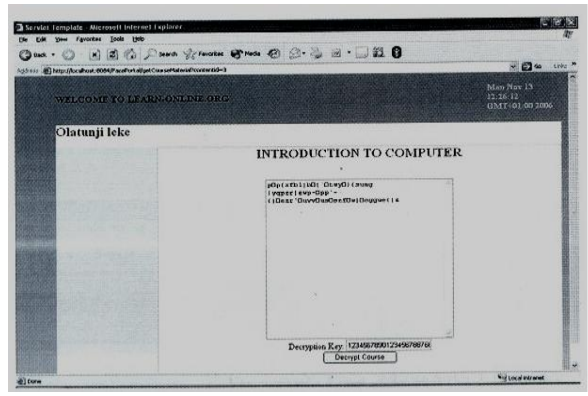

Fig 12 View of Encrypted Course Material

10. CONCLUSION

This study is expected to proffer some solution to insecurity in Computer systems, through development of a crypto-biometric system that implements an improved multilevel security model, optimized facial recognition algorithm and improved data encryption algorithm (tagged XOR-RSA algorithms) as solutions to insecurity of data and development of a Web-based learning System as test bed for these solution.

The methodology used involves the development of an online Web-based learning portal by the use of object oriented programming language, Java that is run over an intranet, with implementation of, eigen facial recognition algorithm (PCA) that is optimized by the use of matlab in order to provide better security for the Web-based learning system.

The outcome of this work is the development of a secured Web-based learning crypto-biometric system built on an enhanced security that is customized with a contextualized framework.

REFERENCES

[1] Jian Yang, Zhang D., Frangi A.F., Jing-yu Yang(2004): Two-dimensional PCA: a new approach to appearance-based face representation and recognition.

[2] IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol 26. 131 – 137.

[3] Oweston, R.D. (1997): "The World Wide Web: a Technology to Enhance Teaching and Public Key Cryptosystem" Common ACM vol 21, pp 120-126

[4] Pentland, I A. (1997)``Coding, Analysis, Interpretation, and Recognition of Facial EMC, pp. 518-526.

[5] Pentland A(2004): Looking at People: Sensing for Ubiquitous and Wearable Computing.IEEE Trans. Pattern Analysis and Machine Intelligence. Vol. 22. 107-119.

[6] Turk and Pentland (1991). Face recognition using eigenfaces. In *Proc. of Computer Vision and Pattern Recognition*, pages 586–591. IEEE, June 1991b.

[7] Turk, M.A., Pentland A.P, (2005): "Face Recognition Using Eigenfaces", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, USA, pp. 586-591 ://www.face-rec.org/algorithms/U.S. GAO. pp. 25-29

AUTHORS' PROFILE

Adagunodo E.R is a Professor of Computer Science of Obafemi Awolowo Uinversity,Ile Ife Nigeria. He lectures at the Department of Computer Science and Engineering in the same University. His research areas are Software Engineering, E-Learning, and Computer Security.

Afolabi A.O is a Senior Lecturer in the Department of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso .Nigeria. Holds P.hD Degree in Computer Science, His research areas are E-learning, Mobile–Learning, Data Security, Biometrics and Software Engineering.