An Improved Authentication and Monitoring System for E-Learning Examination Using Supervised Machine Learning Algorithms

[1]Adetoba, B.T, [2]Awodele, O, [3]Kuyoro S.O, [4]Nwaocha O.V
[1,2,3,4] Department of Computer Science, Babcock University Ilishan, Ogun State, Nigeria
[4] Department of Computer Science, National Open University, Lagos State, Nigeria

Abstract— The development in Information and Communication Technology and the introduction of Internet enabled examinations, have generated security and reliability issues in e-learning. A lot of academic dishonesty, fraud and identity theft have been constantly reported in literature. Over the years, biometric authentication was used but this was susceptible to fake biometric input and biometric-database modification. Hence, the need to develop a model that will curb the vulnerabilities of e-examination in e-learning environment. This work combined password, biometric authentication (facial detection, recognition and verification) and video monitoring using Supervised Machine Learning Algorithms (SMLA) to address the vulnerabilities of e-examinations. The developed model was tested, using 250 facial images (dataset) acquired from the entire National Diploma students of Computer Science, Yaba College of Technology. For model testing, the Mean Square Error (MSE) and the Root Mean Square Error (RMSE) were calculated to determine the efficiency and validity of the model. A window-based application called E-Learning Authentication and Monitoring System (ELAMS) was developed using Java programming language, PHP, JavaScript, jQuery, CSS and HTML. MySQL was used as the database on Apache Server. The result from system testing showed that MSE and RMSE had values of 1.35 and 1.17 respectively, indicating that the solution was efficient and valid. This implies that it was near impossible for any examination fraudster to match the identity of student in the database. This study provided a new methodology for unbroken e-examination authentication and monitoring system with high reliability.

Index Terms— Authentication and Monitoring, Biometric authentication, E-examination, E-learning, Examination malpractice, Machine learning, Supervised Learning.

1.0 INTRODUCTION

In many e-learning systems, user authentication has been a major concern due to technological advancement and the growth experienced in the use of the internet. In the traditional learning environment, candidate can be recognized and observed. However, this may be difficult in a virtual learning environment. More recently, numerous organisations and institutions have embraced Electronic Learning (EL) or online learning as a way of making learning or education accessible to every individual. Over the decades, the population of students that participate in electronic learning has significantly increased [1], this signifies that electronic learning is becoming a popular replacement for the conventional (that is, old or traditional style) classroom setting [2]. Several higher institutions in developing countries such as Nigeria have also ventured into e-learning projects so as to increase efficiency and to take full benefits of its advantages [3]. However, the electronic or virtual learning environment is not without its own drawbacks. Studies have shown that it is difficult to validate or authenticate users online and to determine if or not an assignment was submitted and completed by a valid person or there was cheating. The processes involved in e-learning have led to an increased academic dishonesty or deception [4].

[5] Conducted a survey an online examination fraud and cheating, and revealed that students are liable to cheating behaviours when there was no stringent test-taking policy in place and it is quite easy to cheat in an online environment. Some studies also confirms that cheating is common in online examination or test than the conventional methods because in online or e-learning students can receive assistance or help without the instructor knowing [6]. Similarly, online examination malpractice could range from looking up answers in printed or handwritten materials availably kept for the purpose of cheating, spying the internet through search engines for possible answers and solution to examination problems, to student impersonation during an online examination [7]. It is also worthy to note that the aforementioned means of cheating in an examination is already been practised by student sitting for traditional examination, which examiners are having difficulty controlling. For online examination to be considered valid and reliable, it is important to apply suitable measures that will guarantee a fair test or examination. These measures may include ensuring that a valid or genuine person is taking the test, the e-examination is free from cheating and that candidate is in the right or regulated place (to deter electronic corruption and illegal assistance), confidentiality, client and server software in addition to privacy are secured and proper access is granted to client before submission [8].

One possible way to curb online examination malpractice is to

establish examinations test centres. This could prove very useful. However, in an environment where the students are geographically dispersed, such that few virtual students exist per geographical location, this method could increase the cost both for the institutions and students. Hence, a more effective and applicable method that supports and encourages virtual learning environment is necessary. This work aims at developing an improved e-Learning Authentication and Monitoring System (ELAMS) using logistic Regression (LR) based on Logistic Loss Function (LLF) and Stochiastic Gradient Descent (SGD) to address the vulnerabilities of e-examination. The system combines Biometric authentication, surveillance (webcam) and machine learning techniques to detect academic dishonesty in an e-learning environment.

2.0 RELATED WORKS

Over time, several authentication models have been developed to provide safe and reliable authentication means in e-learning environment. Some authors considered securing both the IT infrastructure and the e-learning application by using biometric authentication approach. [9] Defines biometrics as "the use of recognition method, such as a fingerprint or retina scan, to verify that an individual in front of a computer screen is indeed the person expected". This could be one of the biometric features (unimodal biometric system) or combining two or more biometric features [10]; [11]. Many studies have shown that using a single biometric factor alone is not enough for a secure authentication method [12]; [13] compared to using multiple authentication method [14]; [15].

Multi-modal biometrics combines two or more biometric verification modalities for identification purpose. For instance, a system may combine face and keystroke or face and signature. Studies have shown that multimodal biometric system is a solution to improve reliability and accuracy of biometrics systems [16]. This is because it improves the recognition effectiveness and also provide higher level of security. [17] proposed a bimodal authentication system using fingerprint and mouse movement. The system enhances the security level but was also found to be susceptible to impersonation attacks and illegal assistance. [18] recommended a multi-biometric authentication scheme using fingerprint methods and face recognition techniques. The system re-authenticates when typical behaviours are exhibited. It was observed that this method enhances security level. The use of fingerprint was considered a robust way of identifying human beings but the system was found to be susceptible to impersonation attacks and illegal assistance because monitoring of the e-learning environment was not considered.

[19] proposed a secured e-examination system for an unsupervised examination environment based on two examination models: Integrated and Secure Electronic Examination Unit (ISEEU) and the Smart Approach for Bimodal Biometrics Authentication in Home examination (SABBAH) to ensure a candidate is the correct student, detect cheating actions and apply penalties throughout his examination (See Figures 1 and 2). The model aims at providing an authentication approach to e-learning systems that ensures cheating-free summative electronic evaluation. This uses a mixture of video surveillance and a bimodal authentication method. The model is highly dependent on psychological factors that helps to deter a candidate from cheating but did not take into consideration monitoring of student throughout the examination process or utilised any machine learning algorithms to detect examination fraud.

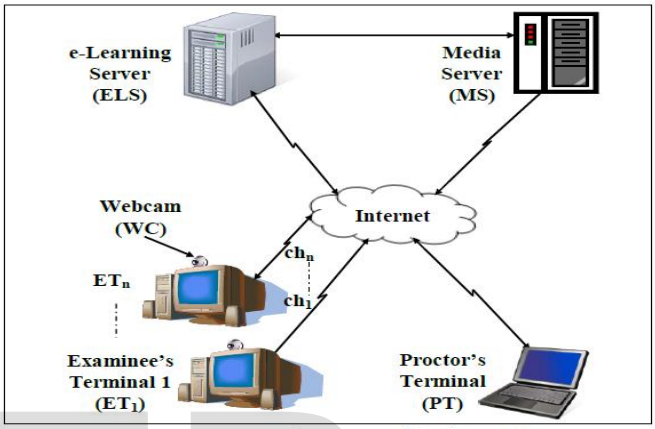


Figure 1: ISEEU model using a webcam [19].

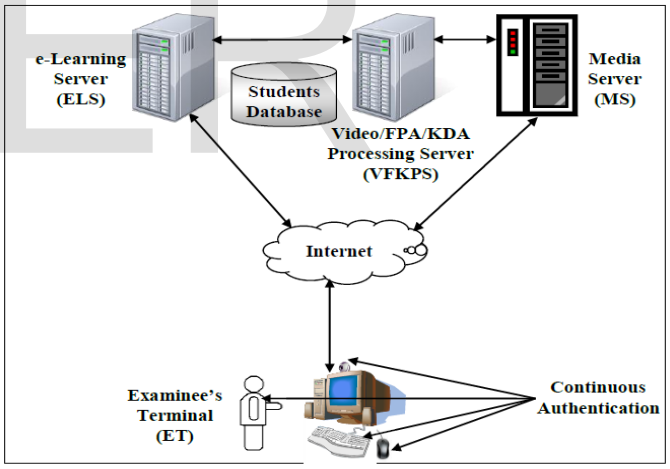


Figure 2: Structure of SABBAH e-examination model [19].

[20] carried out an experiment using five biometric features which were ear, fingerprint, palm print, iris, and retina for identity verification. The result from the experiment shows that biometric multimodal verification method is more reliable, efficient and effective than using single biometric approaches. The biometric features combined were too many and this will increase the cost of implementation. [21] designed an innovative multimodal authentication approach combining at score level, iris and fingerprint traits. Experimental results using fused performance (fingerprint + iris) shows a significant improvement for identity verification compared to using single biometric. This model solves imper-

sonation problems but does not consider monitoring the examination process. [22] carried out an investigation on the role of Multi-biometric Authentication in Professional Certification E-exams. It was concluded in some studies that "using multiple authentication methods that rely on 'who you are' is more secure than using an authentication method that relies on a single 'what you own or know' authentication mechanism" [15]; [23].

[24] reviewed user authentication in e-learning environment. This study shows a complete authentication method that has been used in e-examination, the likely threats that are common during e-exam sessions and the existing commercial user authentication products that are used to observe the e-examination. Some of these authentication methods are based on username and password and this has been discovered to be more vulnerable to several security risks [25]. Some consider using more than one method of authentication to secure electronic examination. [26] build a Profile-Based Authentication Framework (PBAF) for student identification and authentication in e-learning examination. Even though the result shows a positive feedback but the number of questions for continuous authentication that was presented during the examination process were reported to be too many and causes a lot of distractions. [27] developed an e-examination for Distance Learning (DL) which makes use of intelligent voice or speech recognition system. It was concluded that voice-based e-examination system will be of immense benefit to the visually impaired students and also supplement the existing web-based method.

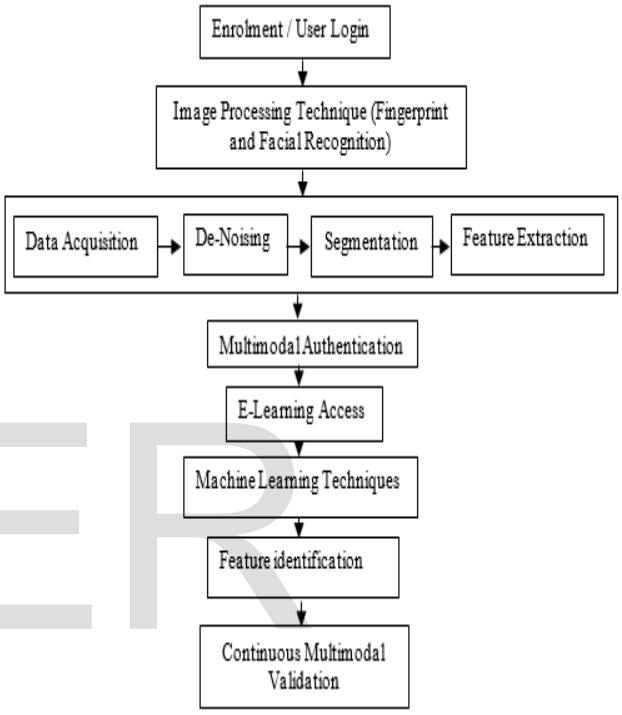
Despite all effort taken to protect the integrity of the e-learning environment, studies have shown that most of the earlier models did not take into consideration combining authentication with monitoring nor fully considered monitoring the e-learning environment throughout the e-examination process. Many studies have centred more on authentication; hence this work focuses more on monitoring using Logistic Loss Function (LLF) and Stochastic Gradient Descent (SGD) in order to detect impersonation in an e-learning setting.

3.0 METHODOLOGY

This system was designed as an improvement on Smart Approach for Bimodal Biometrics Authentication in Home exams (SABBAH) and Integrated and Secure Electronic Examination Unit (ISEEU) models (Figure 1 and 2). Data were collected from students of Yaba College of Technology using purposeful and random sampling from participants who registered on the e-learning portal (both textual and biometric data). The textual data collected include biodata such as Name, Age, Sex, and so on. The biometric data included fingerprint and facial image captured during enrolment. Facial image was captured using the system webcam while a fingerprint scanner is required for capturing digital image of fingerprint as input into the system. These data were collected during registration and serve as training data for both authentication and monitoring. Figure 3 shows the block chart of ELAM phases.

Figure 3: Block chart of the E-learning Authentication and Monitoring System (ELAMS)

The model was designed using multimodal (that is, face and fingerprint recognition system) biometric authentication system. The captured images from the multimodal biometric



undergo pre-processing after which noise is removed by utilizing Gaussian filter and nonlinear filter. The pre-processed images were segmented by separating them into clusters using k-means algorithms. Image processing and features extraction of the pre-processed images segmented were done using the Principal Component analysis (PCA). There was comparison of features stored in the database with each classified data of the same type as well as any abnormality recorded.

To develop the application called eLAMS, Apache was used as the server, MYSQL for the database, PHP as the scripting language and Windows was used as the operating system. The interfaces were connected to the database of pre-processed multimodal fingerprints and facial recognition. Then, the development of pattern recognition algorithms using Logistic Loss Function (LLF) and Stochastic Gradient Descent (SGD) to monitor the pre-processed facial images captured in the database against the ones collected during the examination. The system was developed using JAVA, PHP, HTML Javascript and JQuery

3.1 Logistic Regression (LR)

LR is a function classification algorithm that uses a class for building model and employs a single multinomial logistic regression model in relation to a single estimator. The logistic regression algorithm was chosen for this work because it is not affected by the classification problems of observations that can only assume values in the range of 0 and 1. As an alternative, in approximating the 0 and 1 values immediately, LR forms a one-dimensional model in relation to a converted target variable. In a situation where there are only two classes. Logistic regression substitutes the original target variable.

The logistic regression machine learning algorithm used for the study consists of a loss function and an optimization technique.

i. Logistic Loss Function (LLF)

Loss function (LF) is useful in measuring a good classifier. The difference between prediction *f(x)* and the actual (*yi)* in testing set is measured using a loss function. Loss function is large when what is predicted is larger than the actual label and vice versa [28].

The loss is the cost incurred (that is penalty) when what is determined or expected of the target provided by the Machine Learning model is not exactly equal to the target. The loss function quantifies this cost (penalty) as a single value while the loss is minimized by an optimization technique. Therefore, the logistic loss function ensures that the loss in the face capturing and face matching is not more than 15% because the threshold was set to be 85%.

ii Stochastic Gradient Descent (SGD)

Few samples of facial image of a particular candidate were randomly selected instead of using the whole dataset for the iteration. The batch approach which denotes the total number of samples was used for determining the gradient for each iteration and this batch was taken to represent the whole dataset. However, the whole dataset utilization is greatly useful when considering the minima in a less noisy or random number. When the dataset increases, the computational time increases. This problem was addressed by using a single sample to perform each iteration for face recognition and face verification that involves the larger number of candidates. This sample for performing the iteration is randomly shuffled and selected.

3.2 Development of the ELAMS

The architecture of the system is made up of five major phases which are also segmented into modules. Figure 4 shows the five phases of the architecture. The phases are the Registration/Enrolment Phase, Multimodal Authentication Phase, Multimedia User Monitoring Phase, E-learning phase and the Machine learning Phase.

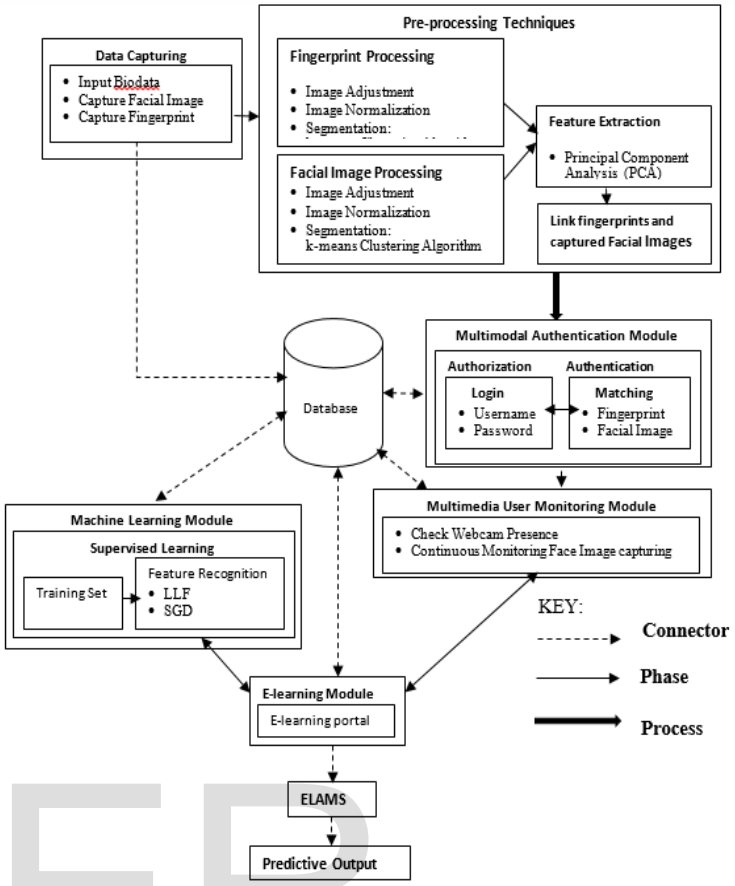


Figure 4: Architectural Model for E-learning Authentication and Monitoring System (ELAMS)

4.0 RESULT

ELAMS was designed such that the student database is accessed once the student has been enrolled, registered and already has a username and password created by the administrator. The system prompts for login details for authentication after which the fingerprint and face recognition will be carried out. For the fingerprint recognition, a standard SDK for Digital Persona was installed on the system. The verification part was done using the Feature-Fusion Extraction algorithm which matches the stored data with the new data capture for a particular user.

For face detection, matching, recognition and verification, an AWS – API (Amazon Web Services - Application Program Interface) called Face recognition was implemented along with the code for faster facial analysis and evaluation considering the system configuration, web services and limitation that could be encountered especially when dealing with face. Therefore, threshold of a confidence score ranges from 0% to 100%. The closer the confidence score to 100, the more indication of the probability that a given prediction is correct. The similarity threshold used for this study for face recognition, matching and verification was 85%.

The developed authentication and monitoring system for e-learning environment has five major modules which comprises of the Registration/Enrolment Phase, Multimodal Authentication Phase, Multimedia User Monitoring Phase, E-learning phase, Machine learning Phase.

4.1 Student Registration and Face Capturing Stage

The enrolment includes: first name, Surname, Email, Phone number, Contact Address, Student Number / Matric Number, Username and Password as well as Picture / Face Capturing which are stored into the database as shown in Figure 5.

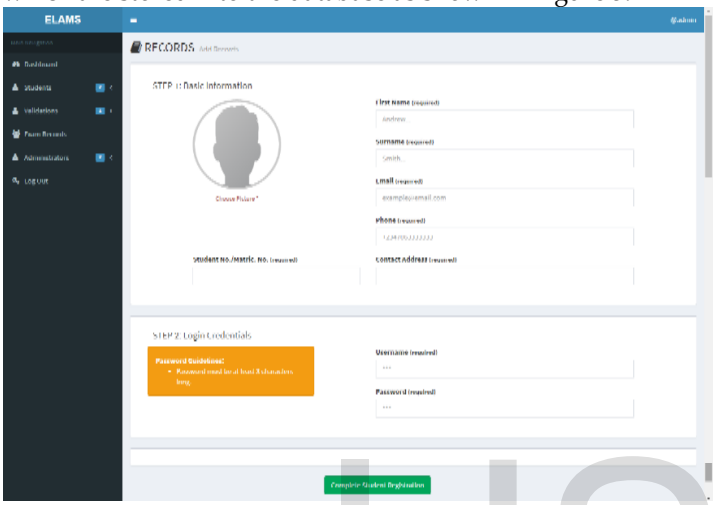


Figure 5: Student Registration and Face Capturing Window by the Administrator

After a successful authorization the student proceeds to the next phase of the ELAMS for face capturing and authentication.

4.2 Face Detection, Matching and Recognition

The system validates the student to access the e-examination portal by first re-capturing student's face after successful login credential and fingerprint have been authorized. This is to ensure that the student whose data have been captured and stored into the database is actually the one sitting for the examination. This is done with the use of logistic regression machine learning algorithm which consists of a loss function and an optimization technique. The optimization technique tries to find how to minimize the loss as small as possible. Figure 6 shows face capturing interface for webcam stream and webcam photo in order to take photo and validate user.

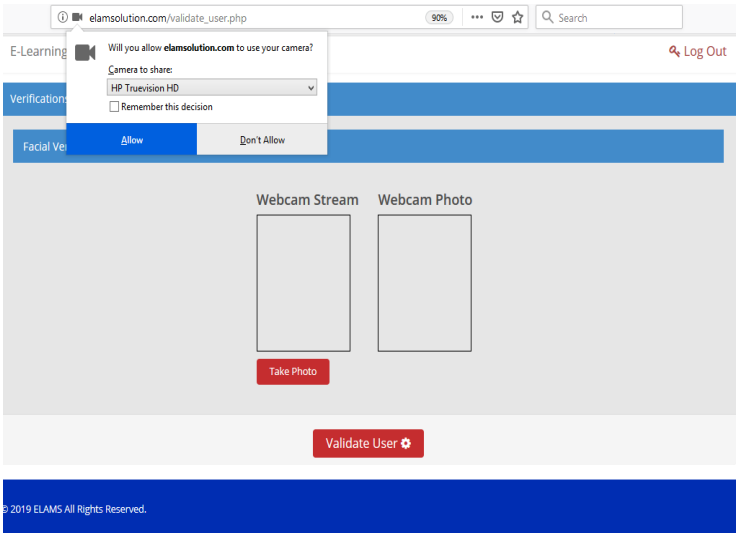


Figure 6: Online Face Capturing

Figure 7 shows successful online face matching and face recognition verification after successful face detection and capturing have been carried out. If the face capturing is not matched successfully, the user will not be authorized to the e-examination module.

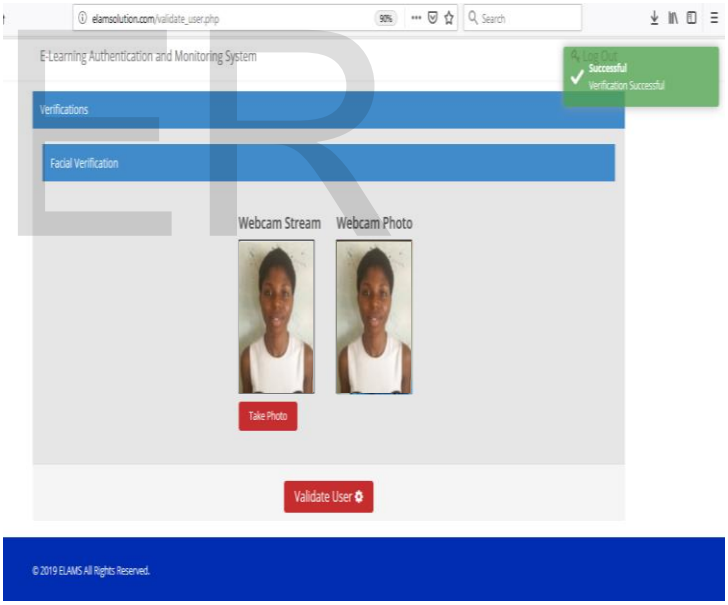


Figure 7: Online Face Detection and Matching

4.3 E-examination and Monitoring

The examination interface is designed with randomized questions so that student performance without any malpractice is assured. Figure 8 show the e-examination environment before the student start the examination.

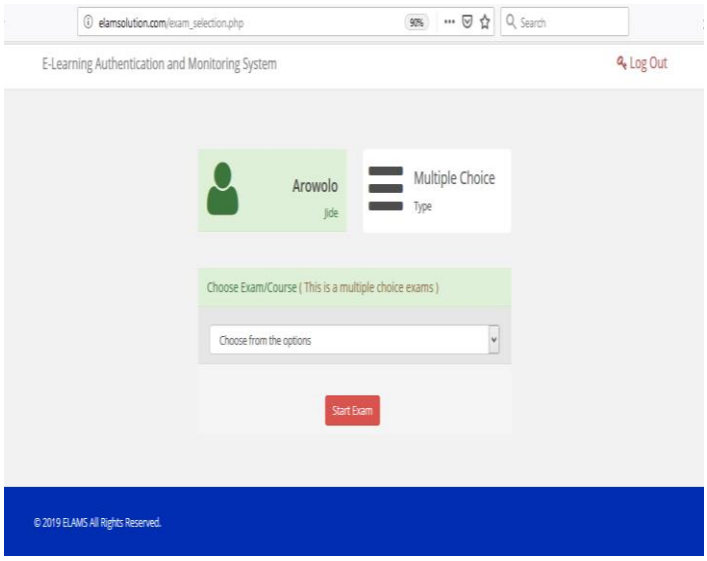


Figure 8: E-examination Environment

Figure 9 shows the interface after the student has clicked on start exam button. Again, the system requested to allow the use of camera. The module once activated will begin the e-examination monitoring process. Again, the webcam will be permanently on until the student logs out or the examination period has elapsed. If the student declines, the examination interface for questions will not load, and the student will be forced to logout. Figure 10 show the interface for e-examination monitoring with the webcam on.

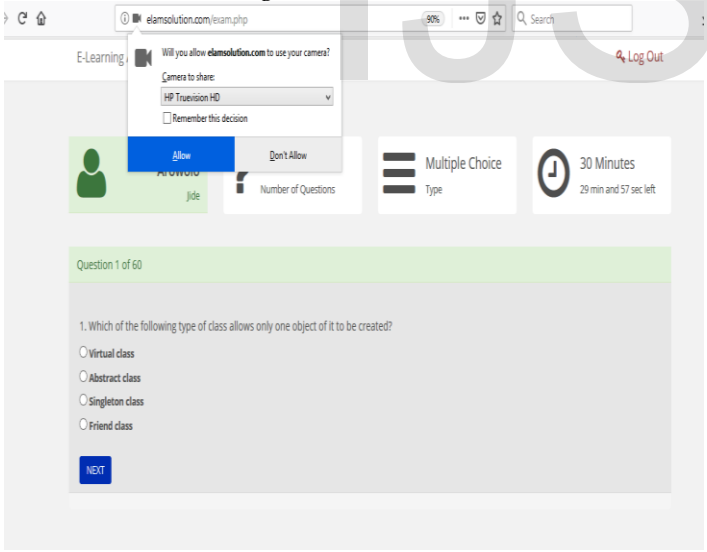


Figure 9: E-examination Environment Verification Interface

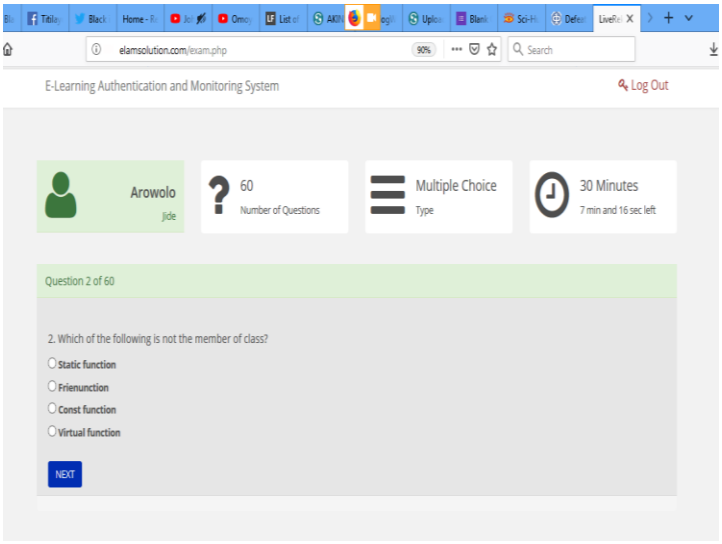


Figure 10: E-examination Monitoring

If there is any impersonation during the examination the ELAMS captures that and the invalid face is saved into the database with status report accordingly. Figure 11 show face validation interface.

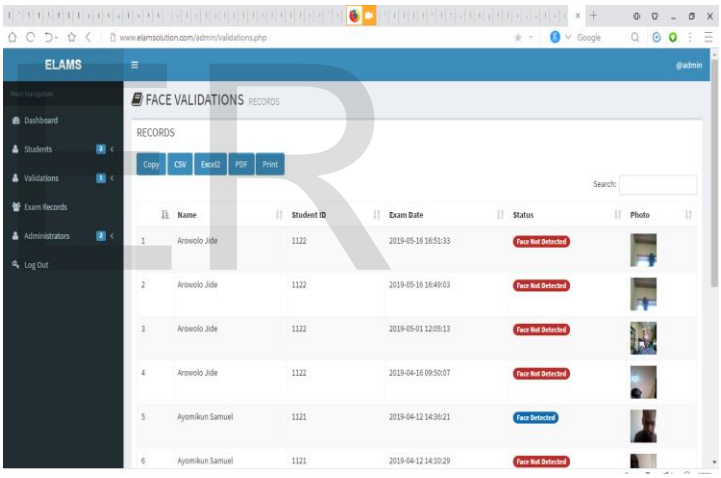


Figure 11: Face Validation Report for e-Examination Monitoring

Figure 12 show examination results reporting with student name, student ID, Registration ID, course, Score and examination date.

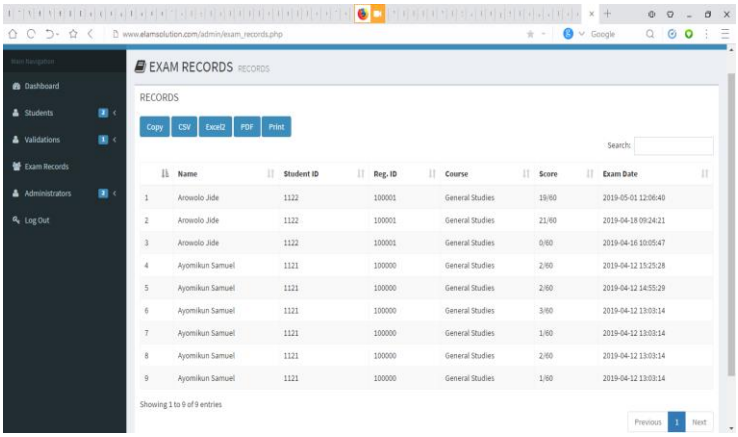


Figure 12: Report for e-Examination Records

5.0 EVALUATION AND TESTING OF THE DEVELOPED SYSTEM

Evaluation and testing of the developed system was done showing Accuracy, Sensitivity and Specificity values of 100% for chosen algorithms namely Logistic Regression and Stochastic Gradient Descent as well as Mean Absolute Error (MAE) values of zero respectively. The study also, revealed that MSE and RMSE have values of 1.35 and 1.162 respectively, which means the solution is efficient and valid. Furthermore, the developed System, ELAMS was subjected to typical indicator for validating the efficiency, performance and accuracy of facial pattern matching as also used [29]. The FAR was 0% and FRR was 1.2%. This clearly shows that, contrary to the claims of the antagonists who reject e-learning that e-examination can never be conducted without malpractices, and that it is devoid of trust and cheating-free, E-learning and E-examination can be trusted. Hence, the study developed a new methodology for unbroken e-examination authentication and monitoring system. This will empower institutions of higher learning to carry out secure e-examinations remotely at any location of the candidates' choice.

6.0 CONCLUSION

It was established that the ELAMS was able to authenticate and closely monitor e-examination environment. The minimization of loss function of logistic regression has played a prominent orole in producing optimal and faster results for accurate predictions especially for face detection, matching, face recognition and verification. Therefore, the study concludes that Logistic Regression can be used of binary image classification and Stochastic Gradient Descent for image optimization for training logistic loss function model for continuous facial image detection, recognition, verification, monitoring and matching processes and researches. The study has shown clearly that security is pivotal to e-learning systems, because of the mode of operating at this present time and the enclosure of sensitive information and operations. E-examination operation has been given so much attention lately, so there is always lack of trust with e-examination for any institutions of higher learning that wants to adopt the approach.

ACKNOWLEDGMENT

We like to acknowledge and appreciate all lecturers and Psotgraduate students of Computer Science Department, Babcock University for their moral support and encouragement during the research. Thank you and God bless.

REFERENCES

[1] I. E. Allen and J. Seaman, "Grade change," Track. Online Educ. United States. Babson Surv. Res. Gr. Quahog Res. Group, LLC, 2014.

[2] R. Pastore and A. Carr-Chellman, "Motivations for residential students to participate in online courses," Q. Rev. distance Educ., vol. 10, no. 3, pp. 263–277, 2009.

[3] A. Fluck, O. S. Adebayo, and S. M. Abdulhamid, "Secure e-examination systems compared: case studies from two countries," J. Inf. Technol. Educ. Innov. Pract., vol. 16, pp. 107–125, 2017.

[4] B. S. Brown and R. Weible, "Changes in academic dishonesty among MIS majors between 1999 and 2004," J. Comput. High. Educ., vol. 18, no. 1, pp. 116–134, 2006.

[5] C. G. King, R. W. Guyette Jr, and C. Piotrowski, "Online exams and cheating: An empirical analysis of business students' views.," J. Educ. Online, vol. 6, no. 1, p. n1, 2009.

[6] M. P. Watters, P. J. Robertson, and R. K. Clark, "Student Perceptions of Cheating in Online Business Courses.," J. Instr. Pedagog., vol. 6, 2011.

[7] C. O. Onyibe, U. U. Uma, and E. Ibina, "Examination Malpractice in Nigeria: Causes and Effects on National Development.," J. Educ. Pract., vol. 6, no. 26, pp. 12–17, 2015.

[8] W. A. Al-Hamdani, "Secure E-Learning and Cryptography," in Cases on Professional Distance Education Degree Programs and Practices: Successes, Challenges, and Issues, IGI Global, 2014, pp. 331–369.

[9] M. Sasikumar, "E-learning: Opportunity and challenges." 2013. Retrieved from http://www.cdacmumbai.in/design/corporate_site/overide/pdf-doc/eLearning.pdf

[10] G. Chandran & Rajesh. "Performance Analysis of Multimodal Biometric System authentication". IJCSNS 290 International Journal of Computer Science and Network Security, 9(3). 2009

[11] A. A. Fathima, S. Vasuhi, T. M. Treesa, N. T. Naresh-Babu, and V. Vaidehi, "Person authentication system with quality analysis of multimodal biometrics," WSEAS Trans. Inf. Sci. Appl., vol. 10, no. 6, pp. 180–194, 2013.

[12] F. Hao, R. Anderson, and J. Daugman, "Combining crypto with biometrics effectively," IEEE Trans. Comput., vol. 55, no. 9, pp. 1081–1088, 2006.

[13] A. Nagar, K. Nandakumar, and A. K. Jain, "A hybrid biometric cryptosystem for securing fingerprint minutiae templates," Pattern Recognit. Lett., vol. 31, no. 8, pp. 733–741, 2010.

[14] H. Ailisto, E. Vildjiounaite, M. Lindholm, S.-M. Mäkelä, and J. Peltola, "Soft biometrics—combining body weight and fat measurements with fingerprint biometrics," Pattern Recognit. Lett., vol. 27, no. 5, pp. 325–334, 2006.

[15] D. Bouchaffra and A. Amira, "Structural hidden Markov models for biometrics: Fusion of face and fingerprint," Pattern Recognit., vol. 41, no. 3, pp. 852–867, 2008.

[16] Y. Faridah, H. Nasir, A. K. Kushsairy, and S. I. Safie, "Multimodal Biometric Algorithm: A Survey," Biotechnology, vol. 15, no. 5, pp. 119–124, 2016.

[17] S. Asha and C. Chellappan, "Authentication of e-learners using multimodal biometric technology," in 2008 International Symposium on Biometrics and Secu-

rity Technologies, 2008, pp. 1–6.

[18] K. M. Apampa, G. Wills, and D. Argles, "An approach to presence verification in summative e-assessment security," in 2010 International Conference on Information Society, 2010, pp. 647–651.

[19] Y. Sabbah, I. Saroit, and A. Kotb, "An interactive and secure e-examination unit (ISEEU)," in 2011 RoEduNet International Conference 10th Edition: Networking in Education and Research, 2011, pp. 1–5.

[20] L. Latha and S. Thangasamy, "Robust Way of Multimodal Biometric Score Normalization," J. Appl. Secur. Res., vol. 7, no. 1, pp. 59–70, 2012.

[21] K. Vishi and S. Y. Yayilgan, "Multimodal biometric authentication using fingerprint and iris recognition in identity management," in 2013 Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2013, pp. 334–341.

[22] G. Smiley, "Investigating the role of multibiometric authentication on professional certification e-examination." Nova Southeastern University, 2013.

[23] A. K. Jain, A. A. Ross, and K. Nandakumar, Introduction to biometrics. Springer Science & Business Media, 2011.

[24] N. A. Karim and Z. Shukur, "Review of user authentication methods in online examination," Asian J. Inf. Technol., vol. 14, no. 5, pp. 166–175, 2015.

[25] S. Alotaibi, "Using biometrics authentication via fingerprint recognition in e-exams in e-learning environment," 2010.

[26] A. Ullah, H. Xiao, M. Lilley, and T. Barker, "Privacy and usability of image and text based challenge questions authentication in online examination," in 2014 International Conference on Education Technologies and Computers (ICETC), 2014, pp. 24–29.

[27] A. A. Azeta, I. A. Inam, and O. Daramola, "Developing e-examination voice interface for visually impaired students in open and distance learning context," in 2017 Conference on Information Communication Technology and Society (ICTAS), 2017, pp. 1–6.

[28] M. Hardt, K. Ligett, and F. McSherry, "A simple and practical algorithm for differentially private data release," in Advances in Neural Information Processing Systems, 2012, pp. 2339–2347.

[29] O. Olatunbosun, "Iwasokun Gabriel Babatunde Akinyokun Oluwole Charles," Perform. Eval., vol. 20, pp. 777–789, 2013.