Multidisciplinary Journal of Management and Social Science (MJMSS)
Volume 1 Number 2, 2024 ISSN: 3043-6427, E-ISSN: 3043-6419
Online publication with Google Scholar indexing, Email: mjmss242@gmail.com

OPTIMIZED FACE AUTHENTICATION FRAMEWORK USING ACTIVE APPEARANCE MODELS

OJUAWO, OLUTAYO OYEWOLE and JIBOKU, FOLAHAN JOSEPH
DEPARTMENT OF COMPUTER SCIENCE, THE FEDERAL POLYTECHNIC,
P.M.B. 50, ILARO, OGUN STATE, NIGERIA
EMAIL: olutayo.ojuawo@federalpolyilaro.edu.ng
folahan.jiboku@federalpolyilaro.edu.ng
PHONE NO: 08038417080, 09062050721

ABSTRACT
In recent years, model-based approaches have garnered significant attention for their effectiveness in image interpretation. These methods offer two primary advantages: first, they ensure robust interpretation by constraining solutions to valid model instances; and second, they enable detailed scene interpretation by explaining an image in terms of a set of model parameters. This approach is particularly well-suited for interpreting facial images, as faces are highly variable, deformable structures that exhibit significant differences in appearance based on factors such as pose, lighting, expression, and individual identity. While model-based techniques have shown considerable success, existing methods often fall short of employing a fully photorealistic model. They typically do not minimize the difference between a model-synthesized instance and the image being interpreted. To address this gap, this study aims to develop an efficient and effective face authentication system using statistical models of shape and appearance through the Active Appearance Model (AAM). This system is designed to process and post private medical reports for doctors as well as patients.The proposed system is designed for applications in facial expression analysis and security authentication. However, when applied to expression recognition or face authentication, these systems tend to exhibit lower accuracy compared to human observers. This limitation suggests potential for further improvement. By integrating the Active Shape Model (ASM) with AAM, researchers can jointly optimize landmark precision and texture mapping, achieving more accurate feature extraction and better matching of image textures. Such a combined approach holds the promise of advancing the field of facial recognition and authentication systems.

KEYWORDS: Face Recognition, Expression Recognition, Active Appearance Model, Authentication.

1. INTRODUCTION
Model-based approaches to image understanding have gained significant interest due to their ability to provide robust and interpretable solutions. These methods offer two main advantages. First, they constrain solutions to represent valid instances of a predefined model, ensuring robust interpretation. Second, they provide a framework to explain images in terms of model parameters, which serves as a foundation for scene analysis and interpretation. These benefits are maximized when the model of object appearance is as comprehensive as possible, capable of generating highly accurate approximations of any target image (Turk & Pentland, 2019).

This paradigm is particularly effective for understanding faces in photographs. Human faces are inherently complex, deformable structures that exhibit significant variability based on pose, lighting, expression, and individual identity. The ability to model and interpret these variations is crucial for extracting valuable information from facial images. Among the various features of facial images, identity is the most frequently sought-after characteristic, making face recognition a key application of such techniques (Ezzat & Poggio, 1996).

Despite the success of model-based techniques, a significant gap remains in the implementation of fully photorealistic models for face interpretation. Existing methods often fail to minimize the discrepancy between a synthesized model instance and the target image under analysis. While satisfactory photorealistic models exist, they often involve a large number of parameters (ranging from 50 to 100) to account for variations caused by individual differences, pose, expression, and lighting conditions. Optimizing such high-dimensional parameter spaces poses a significant computational challenge, limiting the practicality of these models in real-world applications.

Facial analysis as a computer vision task has numerous applications, ranging from identity verification to emotion recognition and human-computer interaction. The success of these applications relies on accurate and efficient interpretation of facial features. This requires a detailed understanding of the data provided by various system components, such as cameras, microphones, and processing units. Accurate data interpretation at this foundational level contributes significantly to the overall performance and reliability of the broader system, particularly in critical scenarios such as security and emergency response.

Several model-based approaches for interpreting facial images have been proposed, each aiming to achieve robust performance by constraining solutions to be meaningful examples of human faces. These models seek to explain the appearance of a given image through a compact set of parameters, which can describe aspects such as pose, mood, and identity. Such capabilities not only improve the accuracy of facial recognition systems but also provide a foundation for a wide array of applications, including medical diagnosis, surveillance, and virtual reality systems.

To effectively interpret a new image, robust techniques are required to identify the optimal match between the image and the model. While correlation-based methods can align a model to an image, they often lack the precision required for high-fidelity applications. For instance, Ezzat and Poggio (1996) developed methods to synthesize new perspectives of a face using a sequence of model viewpoints. Their stochastic optimization approach, although computationally intensive, demonstrated reliability due to the high quality of the synthesized images. However, their work did not propose a practical search strategy for matching the model to a new image.

Additionally, advancements in model-based techniques have introduced three-dimensional representations of facial features, encompassing both shape and appearance. These models consider the complete surface texture and geometry, enabling detailed and realistic synthesis of facial appearances. However, they often face limitations in achieving efficient optimization strategies for matching these models to real-world images. This gap underscores the need for innovative methods to improve the computational efficiency and accuracy of these systems.

Moreover, the application of these techniques extends beyond face recognition to include a wide range of tasks such as gaze tracking, emotion analysis, and augmented reality. These systems process incoming data from user devices such as computers and cameras, filtering and interpreting this information to provide actionable insights. By leveraging robust model-based techniques, these applications enhance the system's overall situational awareness, contributing to its effectiveness in handling complex scenarios.

2.0 LITERATURE REVIEW

In time past, many authors have worked and contributed to the premise of face authentication using active appearance model, in this section, we take a look at some of them.

2.1 ACTIVE SHAPE MODEL (ASM)

The state of an article is addressed through tourist spots which are one chain of successive qualities focuses, every one of which is significant, point existent in the greater part of the pictures being thought of, for instance, the area of the right eye. Enough number of characteristic focuses ought to be given to cover the exhaustive shape and subtleties. In this model, a sum of 68 characteristic focuses are characterized to make sense of the state of a human face, covering the region of the eyebrows, cheeks, eyes, mouth, and nose.
A gathering of tourist spots frames a shape. In the meantime, the shapes are addressed as vectors: every one of the directions sought after by the y-directions of the places in the structure. Adjust one shape to other with a correspondence change (permitting turn, scaling, pivot, and interpretation) that diminish the Euclidean distance normal between shape focuses. The mean shape is pronounced the center of the delineated preparation shapes (Darkner, et. al 2004). The ASM starting the quest for facial milestones from the mean shape adjusted to the spot and size of the face indicated by a worldwide face indicator. It then emphasizes the accompanying two stages until combination.

2.2 DEMONSTRATING FACE APPEARANCE

In this segment, how our appearance models of countenances were created is being framed. The methodology follows that portrayed in Edwards et al (1997) yet incorporates additional dark level standardization steps. Some experience with the fundamental methodology is expected to comprehend the new Active Appearance Model calculation. The models were produced by joining a model of shape variety with a model of the appearance varieties in a shape-standardized outline.

A preparation set of named pictures were required, where milestone focuses are set apart on every model face at key situations to frame the principal highlights. Given such a set, a factual model of shape variety was being created. The marked focuses on a solitary face depict the state of that face. Every one of the arrangements of focuses were adjusted into a typical co-ordinate outline and address each by a vector, x. Then, an essential part investigation (PCA) was applied to the information. Any model can then be approximated utilizing: x = x + Psbs. Where x is the mean shape, Ps is a bunch of symmetrical methods of shape variety and bs is a bunch of shape boundaries. To construct a measurable model of the dark level appearance, every model picture was warpso that its control focuses match the mean shape (utilizing a triangulation calculation). Then inspected the dim level data from the shape-standardized picture over the district covered by the mean shape.
To limit the impact of worldwide lighting variety, this vector was standardized, getting.

2.3 ACTIVE SHAPE MODEL (ASM) ALGORITHM

I. The algorithm of ASM model consists of five steps (Cootes et al, 1994):
II. Labeling the points.
III. Eliciting gray profile to all landmarks.
IV. Aligning training set for ASM.
V. Calculating statistics on aligned training set on PCA at every resolution.
VI. Repeat steps from step 1 to step 4 for all resolution level.

After these processes, for each shape X can be represented as: X = X+ P. b. Where, Xis the ASM mean of training set, eigenvectors P performs the best significant methods of variation, b it is a coefficient vector.

2.4 ACTIVE APPEARANCE MODEL (AAM)

AAM model suggested by Cootes et al. (2004) is one of the strongest model-based objects tracking and detecting algorithms. It is generative, nonlinear, parametric model and can be referred to active contour algorithm “snakes,” (Kass, et. al 1988), and (ASM) algorithm (Cootes, 1999). Especially, the AAM forms the texture, shape of the object to produce a set of immediate and realistic photos. The AAM has been widely applied in different cases (Toth, & Madabhushi, 2012). The most persistent implementation of AAM model has been facial tracking and modelling (Lee, et al, 2009).

ASMs and AAMs vary in the quantity of texture variation they capture. ASMs capture very few texture variations and basically use models of shape, while AAMs include detailed texture (pattern of intensity or color) alteration information are thus completely generative methods that can make photorealistic images. So ASMs match shape models, while AAMs match full models of appearance to an image. In Cootes' report (Cootes et al, 2004) one can detect that Covell explained that the parameters of an Eigen trait model can be applied to drive shape model points to the right place. Cootes' AAM, applied here, is an extension of this idea. AAMs, consider what are valid shapes (valid forms) and intensity variations from a training set, to generate synthetically examples like those in a training-set. The AAM model has a strong modelling and effective ability for fitting the raised complexity due to the high-dimensional texture representation which restricts its application to numerous concepts, such as real time systems.
To prepare the AAM algorithm, which is more suitable to actual applications, more potential should be consumed to optimize the calculation of the AAM.

Therefore, several enhancements are suggested to attain this aim. Henceforth, various methods are suggested to decrease the distance of the texture, for example, wedge let-based regression tree (Darkner et al, 2004) and Haar wavelet (Wolstenholme & Taylor, 1999). Moreover, these processes improve the efficiency of the cost of decreasing accuracy or wasting detail information. A frame process is the inverse compositional image alignment (ICIA) (Baker & Matthews, 2004) algorithm that averts updating texture parameters for each frame and is a fast-fitting algorithm for the AAM model. The determination of this model is that it cannot be used to the active shape model that limit the shape and the appearance variance with a lone group of parameters.

2.5 CLASS-SPECIFIC REFINEMENT OF RECOGNITION FROM SEQUENCE

The imperfections of LDA when applied to a specific individual can be modelled by observing the behavior of the model during a sequence. A class-specific linear correction to the result of the global LDA is described, given a sequence of a face.

To illustrate the problem, a simplified synthetic situation in which appearance is described in some 2-dimensional space is described. Then, many representative training examples for two individuals was imagined, person X and person Y projected into this space. A perfect discriminant analysis of identity would allow two faces of different pose, lighting and expression to be normalized to a reference view, and thus the identity compared. It is clear from the diagram that an orthogonal projection onto the identity subspace is not ideal for either person X or person Y. Given a fully representative set of training images for X and Y, the ideal projection could be worked out in advance. However, it was not wished (or needed) to be restricted to acquiring training data in advance. If an example of person Z was to be identified, for whom there is only one example image, the best estimate possible is the orthogonal projection, A, since it cannot be known from a single example whether Z behaves like X (in which case C would be the correct identity) or like Y (when B would be correct) or indeed, neither. The discriminant analysis produces only a first order approximation to class-specific variation. In this approach, class-specific corrections from image sequences were calculated. The framework used is the Appearance Model, in which faces are represented by a parameter vector c. LDA is applied to obtain a first order global approximation of the linear subspace describing identity, given by an identity vector, d, and the residual linear variation, given by a vector r. A vector of appearance parameters, c can thus be described by: $c = c + Dc + Rr$. Where D and R are matrices of orthogonal eigenvectors describing identity and residual variation respectively. D and R are orthogonal with respect to each other and the dimensions of d and r sum to the dimension of c.

The projection from a vector, b onto d and r is given by d = DT c and r = RT c. This equation gives the orthogonal projection onto the identity subspace, d, the best classification available given a single example. It is assumed that this projection is not ideal, since it is not class specific. Given further examples from a sequence, a class-specific correction is applied to this projection. It is assumed that the correction of identity required has a linear relationship with the residual parameters, but that this relationship is different for everyone.

Formally, if dc is the true projection onto the identity subspace, d is the orthogonal projection, r is the projection onto the residual subspace, and r is the mean of the residual subspace (average lighting, pose, expression) then, d - dc = A(r r), where A is a matrix giving the correction of the identity, given the residual parameters. During a sequence, many examples of the same face are seen. These examples can be used to solve Equation 10 in a least-squares sense for the matrix A, by applying linear regression, thus giving the class-specific correction required for the individual.

3.0 METHODOLOGY

In face recognition application, accurate face alignment has determinative effect. Active Appearance Model (AAM) is one of the most studied methods for accurate locating objects. When applying Active Appearance Model, firstly enough face images were collected with various shapes as training set. Then a set of points was used to annotate face shape, so face shape can be represented by the coordinates of these landmarks. After a series of transformation such as Principal Component Analysis, the mean shape of all the faces can be obtained to construct shape model for face alignment.

Given a new face image, the model’s initial position was estimated, computed the suggested movements, and then got a good face alignment result.
Shape parameters and appearance parameters obtained by alignment for real time video tracking and pose estimate were mainly used. The speed of this algorithm is 123 frames/second. So, it is applicable in real time face tracking.

The pose estimate we implement contains horizontal and vertical movements; forward or backward from camera; rotation angle; eyes and mouth state. Model based algorithm are proved to be effective in image and video processing. This is mainly because these models satisfy the needs for an automatic system to “understand” images, e.g., to recover the image structure. Various algorithms were proposed to fulfill such tasks. Active appearance model (AAM) is one of the most powerful model-based algorithms. It can be traced back to the active contour model and active shape model (ASM). Particularly, AAM decouples and models the shape and the texture of the deformable object and is able to generate a variety of instances photo realistically.

3.1 OVERVIEW OF THE SYSTEM DESIGN

The development of this system is modelled for doctors and patients for medical report purposes. The system uses OpenCV, a library used to capture images or videos on the web and prepares the images for any operation. In this place, it prepares it for machine learning analysis. When a user opens the software, the first thing is to either register or login.

For a new patient, since medical reports needs secrecy, he registers with his details such as name, blood group, genotype and so on, and then captures his image the first time, then the system tells the patient to change position and capture his image the second time for authentication. This image will be processed by the model, the model built with (AAM) will capture the physical appearance of the image (the user's face), this appearance includes the contour, shape, color, age, size etc. of the face. An image ID will be created for the user's face authentication. Along with the face attributes, face id, the user's email will be saved together in the model, and the model serves as the database here too. To login, the user needs to enter his email and then scan his face with the application. The model compares the user's active appearances with the one saved, if they match at a relevant level, then the user can gain access to the system, afterwards, the newly scanned face will be added to the training model, so the system can easily identify and recognize the user's face for later use.

3.2 INPUT DESIGN

Input is the data that should be fed into the system which serves as basis for the desired output. The input to the new system is designed to capture data in the subsystem.

LOGIN

EMAIL

Click here to scan your face and login

Figure 1. **Input Design**

3.3 OUTPUT DESIGN

With the identification of the input design, it is logical to state the output design too. Output is the data that should be displayed on the system when the input data has been provided. The output design of this system includes reports that can be viewed by the students, lecturers and the administrator.

LOGIN SUCCESSFUL

FACE VERIFIED SUCESSFULLY

Figure 2. **Output Design**

3.4 FLOWCHART OF THE SYSTEM

Flowcharts are graphical representation of steps. It was originated from computer science as a tool for representing algorithms and programming logic, but had extended to use in all other kinds of processes. Nowadays, flowcharts play an extremely important role in displaying information and assisting reasoning. They help us visualize complex processes, or make explicit the structure of problems and tasks

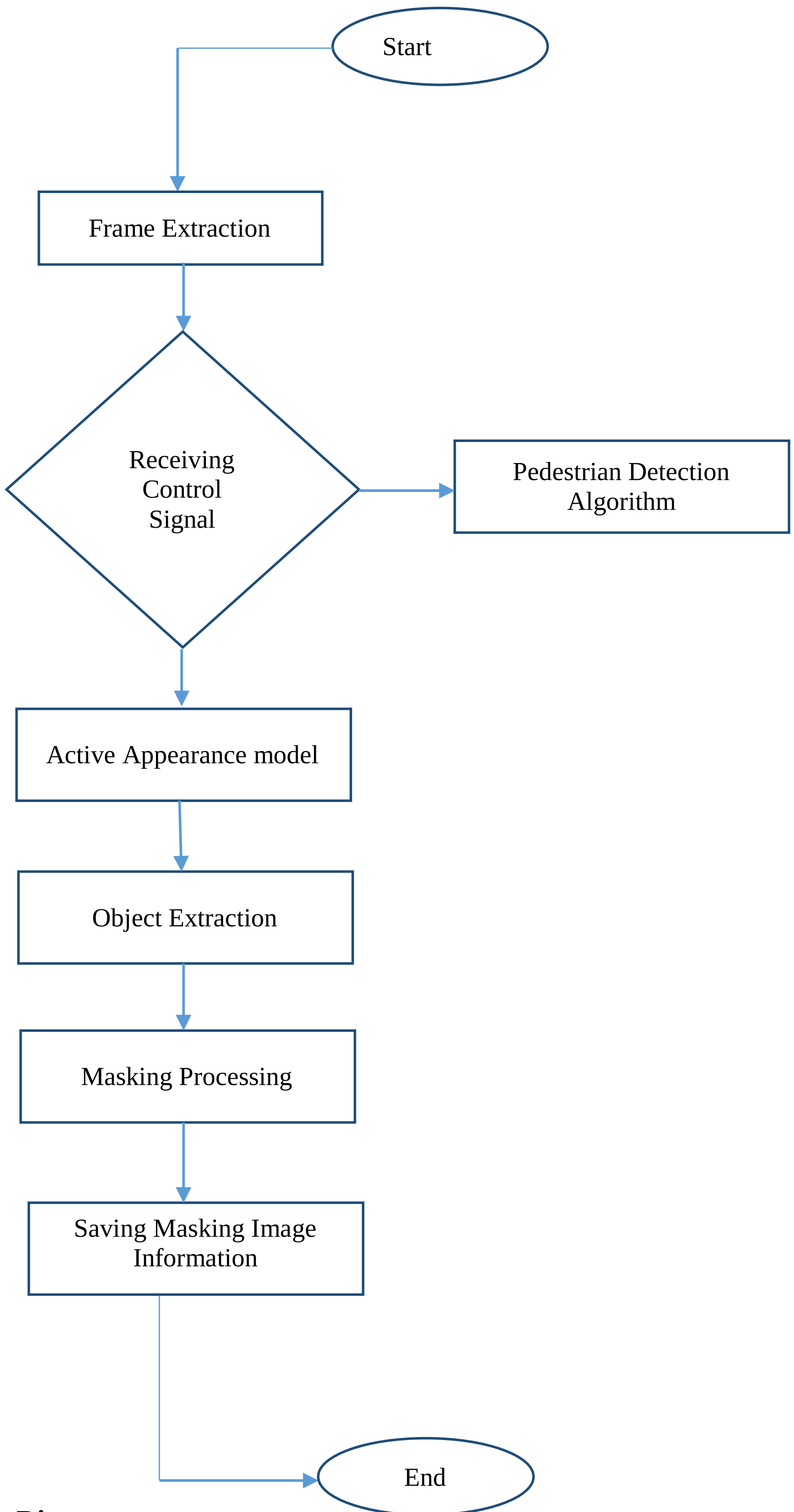

Figure 3. **Flowchart Diagram**
Source: Compiled by Author

Multidisciplinary Journal of Management and Social Science (MJMSS)
Volume 1 Number 2, 2024 ISSN: 3043-6427, E-ISSN: 3043-6419
Online publication with Google Scholar indexing, Email: mjmss242@gmail.com

3.5 SYSTEM ARCHITECTURE

The architecture of this system specifies the hardware, software, access methods and protocols used throughout the system. It also shows the interconnectivity among the components and how they work together to achieve a successful system.

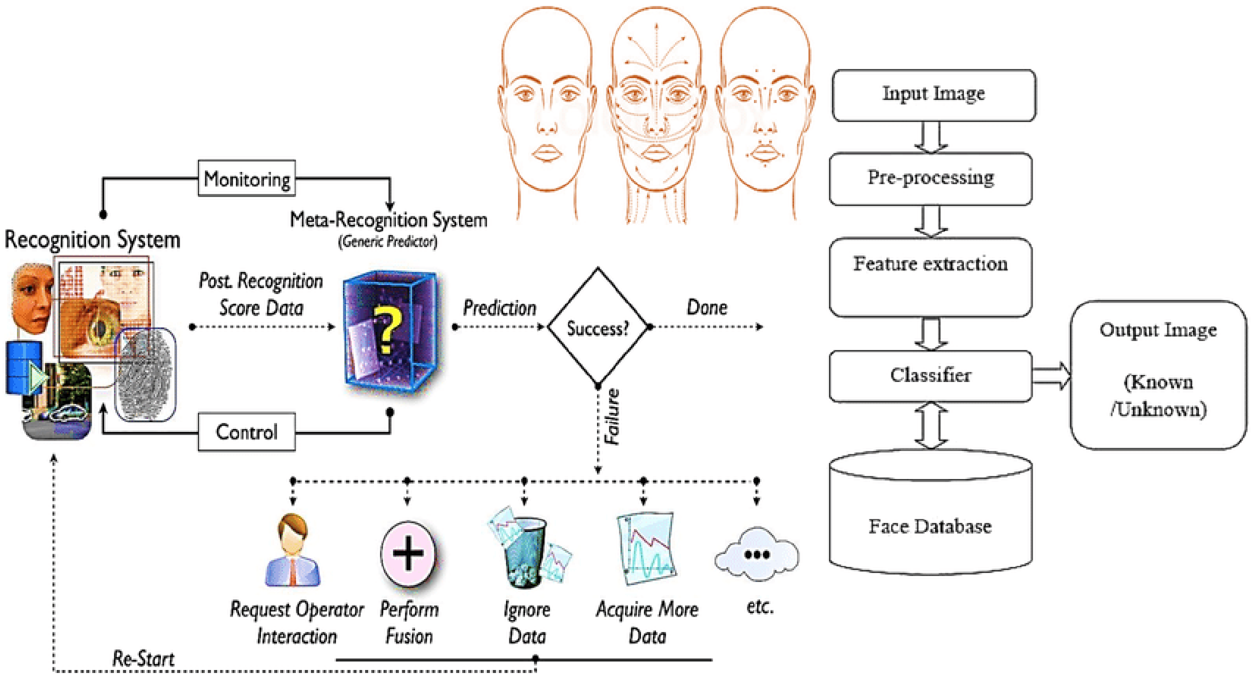

Figure 4. **Architecture diagram**

4.0 RESULTS AND DISCUSSION

The input of the new system contains complete and concise basis for face authentication system using active appearance model while the output contains the relevant information needed by both the users of the system.

4.1 SYSTEM IMPLEMENTATION

Systems implementation is a set of procedures performed to complete the design (as necessary) contained in the approved systems design document and to test, install, and begin to use the new or revised Information System. For the implementation of the new system, data must be prepared for live testing. The result from the new system is compared to that of the existing system to check if the expected result was achieved. It is also necessary to formulate the operation of the new system to check the overall time and ability of the deposit to be handle by the new system.

4.2 DISPLAY OF GRAPHICAL USER INTERFACE

The system consists of several screens and components working independently together as one to achieve the aim of the system. The screens are displayed below showing their functions too:

The Landing Page

This page is the first page displayed by the system that is if the user is not yet logged into the system. He or she will either choose to login or register.

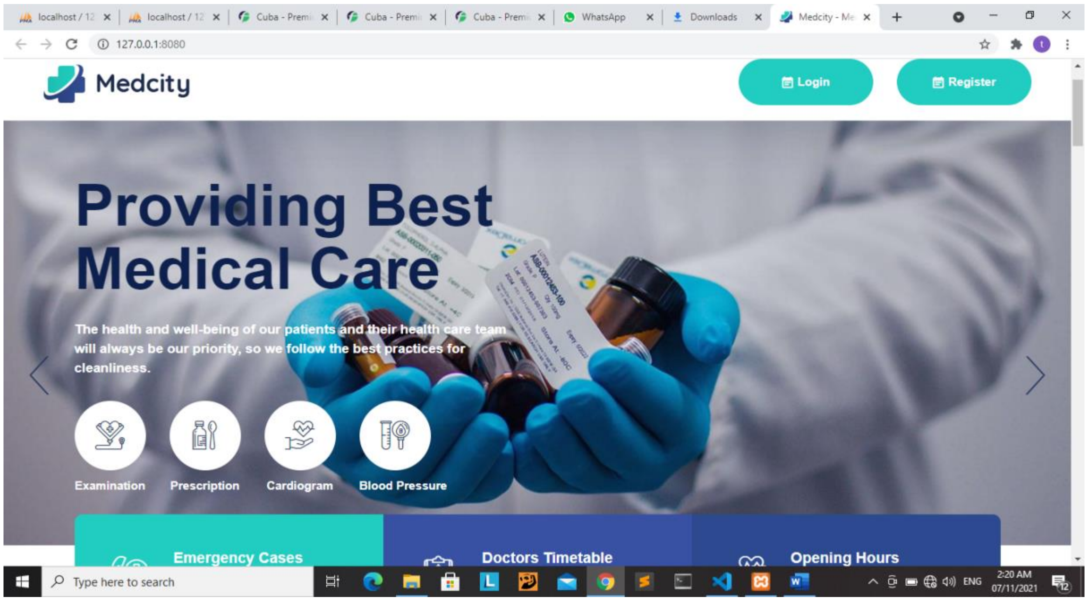

***Figure 5*. The Landing Page**

Register Page

This page allows students to enter their details for registration. Details like username, full name etc. are inputted to allow access into the system.

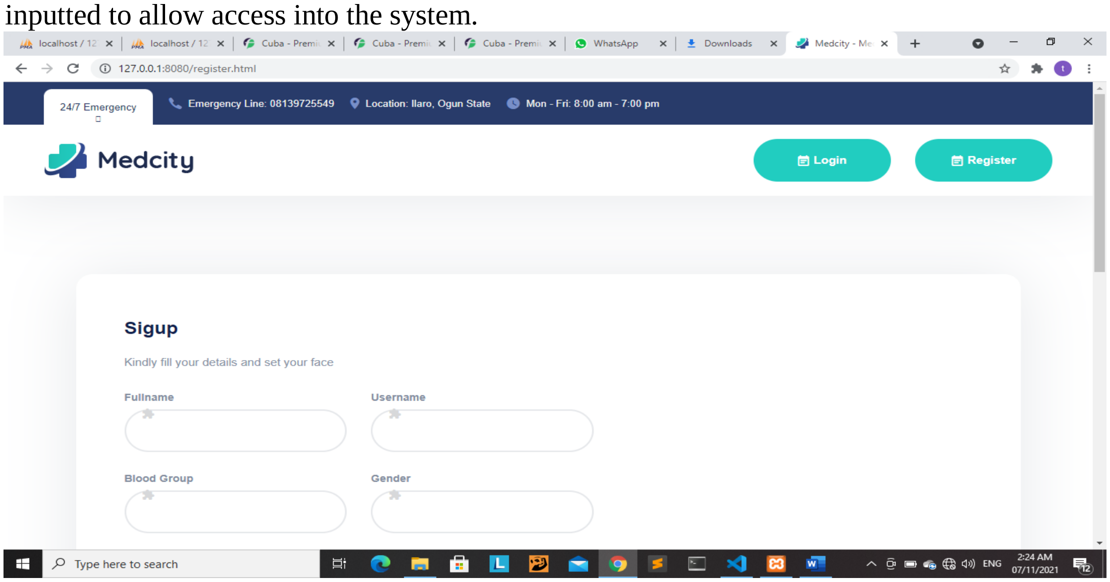

***Figure 6.* The Register Page**

Login Page

This page allows patients to enter their details for authentication. Patients enter their username to access into the system. If the authentication fails due to invalid credentials, the system will prompt the patient to enter a valid one.

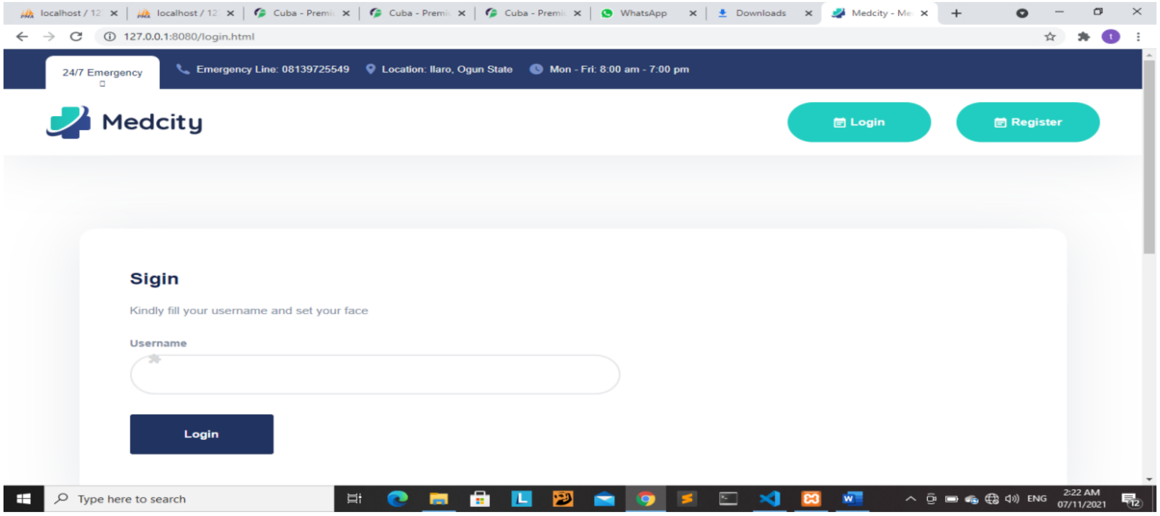

***Figure 7*. The Login Page**

Verification Page

This page shows how verification is done by comparing face with the one provided during registration. If the face matches, then patient is granted access to the system, else it request a valid face.

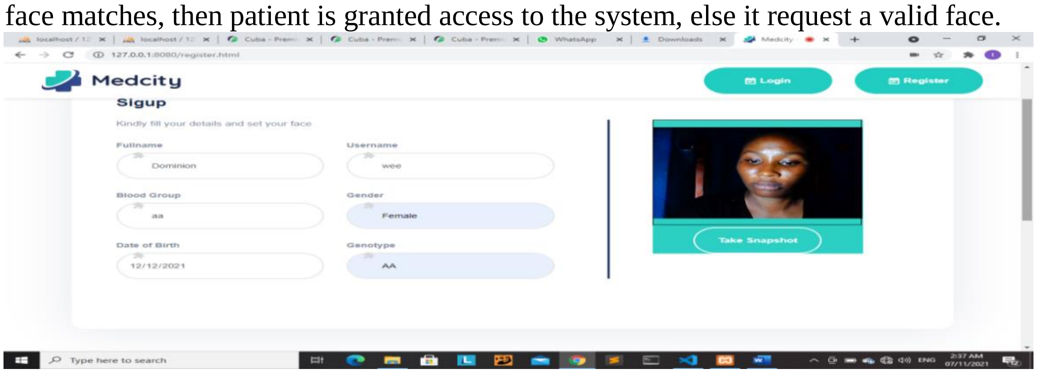

***Figure 8*. The Login Page**

5.0 SUMMARY

This research work is a system that facilitates face authentication system. This system is modeled for doctors and patients for processing and posting of confidential medical reports. This system id developed using the Active Appearance Model. With this system, the doctors can post medical reports and the patient who owns the medical report can only have access to the report through face authentication. Finally, documentation was also provided in case there is need for future changes. Above all, there was also user's document for anyone that wants to use the package.

5.1 CONCLUSION

The use of an Active Appearance Model in face authentication has been described. The model uses all the information available from the training data and facilitates the decoupling of model into ID and non-ID parts. When used for static face identification the AAM proved as reliable as labelling the images by hand. An identification rate of 88% was achieved. When used for expression recognition the systems show less agreement than human observers but nevertheless encourages further work in this area. An observation of the quality of model fit, and the excellent identity recognition performance suggests that the classifier itself rather than the AAM search limits the expression recognition performance.

A technique for improving the stability of face identification and tracking when subject to variation in pose, expression and lighting conditions has been outlined. The tracking technique makes use of the observed effect of these types of variation to provide a better estimate of identity, and thus provides a method of using the extra information available in a sequence to improve classification. By correctly decoupling the individual sources of variation, it is possible to develop decoupled dynamic models for each. The technique described allows the initial approximate decoupling to be updated during a sequence, thus avoiding the need for large numbers of training examples for everyone.

5.2 RECOMMENDATION

It is highly recommended that this research be improved upon and implemented as this will be very useful to the medical field and to the society at large. The system creates a confidential platform for doctors and patients.

In the future work these results can be compared to other Models which with face recognition also the researchers can improve and build a combined model using ASM and AAM model that jointly optimizes a precise traits point location and gives a better match to the image texture.

REFERENCES

Baker, S. B, & Matthews, I. O (2004). Lucas-kanade 20 years on: A unifying framework. *International journal of computer vision*, 56(3), 221-255.

Cootes, T. F., Edwards, G., & Taylor, C. J. (1999). Comparing Active Shape Models with Active Appearance Models. *Procedings of the British Machine Vision Conference 1999*. https://doi.org/10.5244/c.13.18

Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. (2004). *Active Shape Models-Their Training and Application. Computer Vision and Image Understanding*, 61(1), 38–59. https://doi.org/10.1006/cviu.1995.1004

Darkner, S., Larsen, R., Stegmann, M. B., & Ersboll, B. K. (2004). Wedgelet Enhanced Appearance Models. *2004 Conference on Computer Vision and Pattern Recognition Workshop*. https://doi.org/10.1109/cvpr.2004.478

Ezzat. T and T. Poggio. Facial Analysis and Synthesis Using Image-Based Models. In International Workshop on Automatic Face and Gesture Recognition 1996pages 116–121, Killington, Vermont, 1996.

Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1(4), 321–331. https://doi.org/10.1007/bf00133570

Lee, S. J., Park, K. R., & Kim, J. (2009). A comparative study of facial appearance modeling methods for active appearance models. *Pattern Recognition Letters*, *30(14)*, 1335–1346. https://doi.org/10.1016/j.patrec.2009.05.019

Turk. M and A. Pentland. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

Toth, R., & Madabhushi, A(2012).AAM-based palm segmentation in unrestricted backgrounds and various postures for palmprint recognition.*In Pattern Recognition Letters*, 34(9), 955-962.

Wolstenholme, C. B., & Taylor, C. J. (1999). Wavelet compression of active appearance models. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 544-554)*. Springer, Berlin, Heidelberg.