Adaptive Regression Splines Models for Predicting Facial Image Verification and Quality Assessment Scores

A. A. Abayomi-Alli, E. O. Omidiora, S. O. Olabiyisi and J. A. Ojo

***Abstract*— Many biometric applications are faced with enormous performance challenges due to submission of low quality facial images. In this study, adaptive regression splines (ARES) models were built for predicting algorithm matching scores (AMS) and overall quality scores (OQS). A face verification and image quality assessment (FVIQA) framework was adopted to extract five facial quality features from still images. The SCface database was adopted for the training and testing datasets with 2,093 and 897 images respectively. ARES models were built from the normalized individual quality scores and algorithm matching scores using ARESLab in the MATLAB environment. A black face surveillance camera (BFSC) database of 50 subjects was populated to mimic the SCface database and act as the target dataset for the model validation. Results from the study shows that FVIQA quality scores and other experimental results are comparable and consistent with previous research works. The model ANOVA decomposition showed that pose variation is the major determinant for model OQS and AMS with 0.046 and 0.261 respectively. From the performance evaluation, model OQS achieved 99.96% and 99.81% prediction accuracy on the test and target datasets while model AMS achieved 87.04% and 84.73% respectively. Subsequently, no failure-to-acquire (FTA) was recorded when superior face images were selected from the SCface database using the developed image verification and quality assessment (IVQA) number.**

***Index Terms*— Adaptive, Algorithms, Biometrics, Facial Recognition, Image verification, Regression models, Quality.**

I. INTRODUCTION

FACIAL recognition is the identification of humans by the unique characteristics of their faces. It is a vividly researched area of computer vision, pattern recognition and more precisely biometrics [1-4]. It has become an important part of everyday life due to increasing demand in security and law enforcement applications. Face recognition has attracted a lot of attention because of its potential applications. Interest is still on the rise, since it is also seen as an important part of next-generation smart environments [5]. Face recognition systems like all other biometrics applications must cope with real world uncontrolled and dynamic environments [6,7]. Hence a face recognition system (FRS) is plagued by a number of intrinsic and extrinsic variations that directly affect its recognition performance. Variations due to low quality images plaque all biometric systems, such variability is due to a long list of factors which includes facial expressions, illumination conditions, pose, presence or absence of eye glasses and facial hairs, occlusion, and aging [8].

Image quality is a characteristic of an image that measures the perceived image degradation; typically, compared to an ideal or perfect image [9]. Imaging systems may introduce some amounts of distortion or artifacts in the signal, so the quality assessment is an important problem. The primary goal of image quality assessment is to supply the quality metrics that can predict perceived image quality automatically. By defining image quality in terms of a deviation from the ideal situation, quality measures become technical in the sense that they can be objectively determined in terms of deviations from the ideal models. These variations in image quality vary significantly depending on where and when the system operates [10]. Posit that the quality of biometric data is operationally important because it directly influences recognition performance while [11] concluded that a major research area is the study of face recognition over a wide range of quality factors. Although there has been a significant improvement in face recognition performance during the past decade, it is still below acceptable levels for use in many applications [12]. This is because different face recognition algorithms are designed to be robust to particular subsets of these factors. Hence, a high quality image for one algorithm is not necessarily of the same quality for another. Therefore, quality should be learned for a specific face matching algorithm [13].

II. PROBLEM STATEMENT

The Image quality values can be used in different stages of biometric applications, some of these include: enrollment-phase quality assessment, verification/identification quality

A. A. Abayomi-Alli is with the Sound, Systems and Structures Laboratory, University of Pittsburgh, USA. He is on research visit from the Department of Computer Science, Federal University of Agriculture, Abeokuta, Nigeria. (Corresponding Author: +14126141116, +2347030672420, e-mail: abayomiallia@unaab.edu.ng).

E. O. Omidiora is a Professor with the Department of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria (e-mail: eoomidiora@lautech.edu.ng).

S. O. Olabiyisi is an Associate Professor with the Department of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria (e-mail: soolabiyisi@lautech.edu.ng).

J. A. Ojo is a Senior Lecturer and Head of the Electronic and Electrical Engineering Department, Ladoke Akintola University of Technology, Ogbomoso, Nigeria (e-mail: jaojo@lautech.edu.ng).

assessment, prediction of algorithm failure, quality-based adaptation of the processing phase and multimodal biometric fusion [10,14-17]. While steady progress is registered each year in face recognition research, real world deployment of biometric verification systems perform far less than the results obtained in the laboratory. The reason is simple, biometric system performance is directly affected by the quality of the images captured in real world and those present in the database. That is, if the quality of the biometric image is poor, the recognition system's performance is certain to be reduced [11,18]. Thus, knowing the quality score of a probe image that will be predictive of the algorithm's accuracy in matching it to its high quality gallery image will provide information that will improve the overall system's performance. Further results from the research will help to provide feedback that will improve a facial sensor (camera) or facial image collection system's performance. With such knowledge, real-time quality assessment can be done in typical face authentication or verification applications where image samples of very low quality can be recaptured, hence improving the reference database integrity, poor samples can be processed using different algorithms and thresholds, higher quality samples can dominate fusion multi-classifier verification systems, etc.

Although a lot of researchers have proposed and reported results claiming to solve the biometric system performance prediction problem using facial image quality but none has been able to solve the problem efficiently in all settings. The reason is a lot of the proposed techniques in literature used only one property of the face or one feature within the recognition process to assess facial image quality. This is contrary to [19] which concluded that no single quality metric can reliably measure biometric system performance. Secondly, these techniques have proved to be inappropriate for verification scenarios where the performance of a recognition algorithm is a function of the probe image's quality when compared with the gallery image [20]. Multiple facial features has been considered by researchers such as [21] that proposed the first overall quality assessment scheme for facial images based on statistical learning. The results from the research were significantly more consistent with the human perception when compared to previous studies. However, the subjective quality assessment protocol used has been reported as cumbersome and expensive to implement by several researchers including [22] thus obtaining objective quality scores is a preferred option that will eliminate the need for expensive subjective studies [23].

In this paper, we attempt to develop adaptive regression splines models for predicting facial image verification and quality assessment scores from a facial image verification and quality assessment framework (FaceIVQA) that measures five image quality attributes through a full-reference objective quality assessment protocol.

III. RESEARCH METHODOLOGY

A. Research Approach

In this research work, experimental data from [24] for the face verification and image quality assessment (FVIQA) framework was adapted as our model data. FVIQA employs an objective full-reference feature extraction of facial images in database for image quality assessment. The developed scheme was designed to extract the faceness, pose, illumination, contrast, and similarity measures of facial images with reference to their high quality gallery pair. Composite image quality scores was calculated and adaptive regression splines (ARES) models were built to predict recognition and algorithm matching scores using multivariate adaptive regression splines (MARS) technique.

B. Data Acquisition

The primary data used in this work was collected from the verification experiments conducted using the FVIQA framework in [24]; other secondary data sources include the SCface database (training and testing dataset) and BFSC database (target dataset) that was collected specially for this study. The face authentication protocol proposed by [25] was adopted because it models true surveillance scenarios. The day-time and night-time test scenarios were followed strictly resulting in the use of 2,990 images from the surveillance camera (SCface) database by [12]. Each subject's gallery image was compared (verification) with twenty-three images of varying qualities including the twenty-two probe images and the subject's high gallery image. At the end 2,093 training and 897 testing trails were conducted.

The FVIQA framework was developed to combine feature extraction techniques for five facial quality measures such as pose (Q_P), faceness (Q_F), illumination (Q_L), contrast (Q_C) and similarity (Q_S). The approach was aimed at extracting image quality values that are effective and will highly correlate with the recognition matching scores. The framework accepts a low quality probe image from the file, folder or computer's webcam. It compares the probe image with the high quality gallery image and continues with the face recognition steps such as image pre-processing, face detection, feature extraction before entering the face verification and quality assessment part.

C. Experimental Facial Databases

Many publicly available facial databases have been developed but not many of them have sufficient image quality variability and subject size to meet the objectives of this research work. In this work, the surveillance cameras face database (SCface) was used for the training and testing datasets while a black face surveillance cameras database (BFSC) was collected specifically for this research as the target database.

1) Training and testing database

Surveillance camera (SCface) database was used as the primary database.

2) Target database

A black face surveillance camera database (BFSC) was collected specifically for this study as a target database to validate the built ARES models and benchmark the result of the validation with that of the testing dataset. The BFSC database consists of images of fifty (50) black subjects and was designed to mimic the SCface database. The variations between the BFSC and SCface databases are:

a) the BFSC is a black face only database;
b) the quality and types of surveillance cameras differs;
c) the camera-to-eye distance differs;

d) the BFSC has distinct illumination levels with three control lights.

D. Data Pre-processing

Pre-processing was done prior to the ARES model building. The raw data collected from the FVIQA verification experiments cannot be used directly to develop the ARES models because most machine learning algorithms requires that all data should be normalized between the range of zero and one.

1) Score normalization

In order to obtain the most correlated and dissimilar data from the original quality measures (Q_F, Q_P, Q_C, Q_L, Q_S), four data normalization techniques were evaluated namely Min-max, Tanh estimator, Z-score and Decimal scaling. A descriptive statistics and correlation analysis was carried out to determine the most effective and robust normalization technique among the four before the final model data was determined.

2) Overall quality score fusion

An overall-normalized score is obtained by the fusion of the normalized quality scores ($Q^{'}$) using the Sum rule. This is simply the sum of all normalized quality measure scores. Thus a composite score known as the overall quality score(OQS) is derived as:

$$OQS = \sum_{i=1}^{N} Q^{'} \quad (1)$$

This overall quality score (OQS) is expected to be predictive of the contribution of the probe image to the performance of the recognition algorithms used.

E. Building the MARS Models

Adaptive regression splines toolbox (ARESLab) was employed for building the prediction models. ARESLab is a Matlab/Octave toolbox for building piecewise-linear and piecewise-cubic regression models.

To build the ARES model, function *aresbuild* in ARESLab was used. The function *aresbuild* builds a regression model using the multivariate adaptive regression splines (MARS) technique [26]. The input variables for the ARESLab MARS are the normalized values of the faceness quality measure ($Q_F^{'}$), pose quality measure ($Q_P^{'}$), contrast quality measure ($Q_C^{'}$), luminance quality measure ($Q_L^{'}$) and similarity quality measure ($Q_S^{'}$). The outputs of the MARS algorithm are overall quality score(OQS) and algorithm matching score (AMS).

Thus,

$$x = [Q_F^{'}, Q_p^{'}, Q_C^{'}, Q_L^{'}, Q_S^{'}] \text{ and } y = [OQS, AMS] \quad (2)$$

To call the *aresbuild* function:

[*model, time*] = *aresbuild* (*Xtr, Ytr, trainParams, weights, modelOld, verbose*)

Inputs are:

Xtr: a vector of 2,093 training data cases of the independent variables ($Q_F^{'}, Q_p^{'}, Q_C^{'}, Q_L^{'}, Q_S^{'}$).

Ytr: a vector of 2,093 of the training data cases of the dependent variables (OQS, AMS).

Outputs are:

The main output is the built *MARS model* with the following components:

coefs: coefficient of the vector of the regression model.

trainParams: a structure of training parameters for the algorithm.

MSE: mean square error (MSE) of the model in the training data set.

GCV: generalized cross validation of the model in the training data set.

Time: algorithm execution time in seconds (s), e.t.c.

MARSplines constructs the relationship between the dependent and independent variables from a set of coefficients and basis functions that are entirely "driven" from the regression data. The goal is to obtain a useful approximation to each function using a set of training data. It employs expansions in piecewise linear basis functions of the form $[x-t]_+$ and $[t-x]_+$. The + means positive part, stated by [27] and [28] as:

$$[x-t]_+^q = \begin{cases} (x-t)^q & , \quad if\ x > t \\ 0 & , \quad otherwise \end{cases} \quad (3)$$

$$[t-x]_+^q = \begin{cases} (t-x)^q & , \quad if\ x < t \\ 0 & , \quad otherwise \end{cases} \quad (4)$$

MARS uses two steps for improving the model building. Many basis functions were added to get better model performance, however the built MARS can show over fitting problem due to large number of basis functions. Backward algorithm was used to preventing over fitting by deleting redundant basis functions in order to obtain the final model. The generalized cross validation (GCV) is adopted to delete the redundant basis functions [29]. The expression of GCV is given as:

$$GCV = \frac{\frac{1}{N}\sum_{i=1}^{N}\left[y_1 - \hat{f}(x_i)\right]^2}{\left[1-\frac{C(B)}{N}\right]^2} \quad (4)$$

where *N* is the number of data and *C(B)* is a complexity penalty that increases with the number of basis function in the model and which is defined as:

$$C(B) = (B+1) + dB \quad (5)$$

where *d* is a penalty for each basis function included into the model and *B* is number of basis functions in eqn. 3 and 4. The details about *d* are given by [26].

After the forward selection and backward deletion phases of the model building algorithm, the final model for ARES model OQS is given as:

$$OQS = 0.68757 + \sum_{m=1}^{10} a_m B_m\,(x) \quad (6)$$

The basis functions B_m and the corresponding coefficients a_m are given in eqn. 7:

$$\mathrm{OQS} = 0.68757 + 0.2*\mathrm{BF1} - 0.2*\mathrm{BF2} + 0.2*\mathrm{BF3} - 0.2*\mathrm{BF4} + 0.2*\mathrm{BF5} - 0.2*\mathrm{BF6} + 0.2*\mathrm{BF7} - 0.2*\mathrm{BF8} + 0.2*\mathrm{BF9} - 0.2*\mathrm{BF10} \quad (7)$$

The final ARES model AMS is given as:

$$AMS = 0.043249 + \sum_{m=1}^{13} a_m B_m\,(x) \quad (8)$$

while the basis functions B_m and the corresponding coefficients a_m are in eqn. 9:

$$AMS = 0.043249 + 3.0208 * BF1 - 1.1097 * BF2 - 130.76 * BF3 - 3.1741 * BF4 - 0.098423 * BF5 + 231.5 * BF6 - 7.6407 * BF7 + 100.55 * BF8 + 6.6726 * BF9 - 45.96 * BF10 + 217.9 * BF11 - 13.867 * BF12 + 84.547 * BF13 \quad (9)$$

Other ARESLab functions used for this research are *aresparams* for configuring the ARES model building algorithm; *arespredict* for making predictions using an the ARES model; *arestest* for testing the ARES model on a test dataset; *arescv* tests the ARES performance using k-fold cross validation; *aresanova* for performing ANOVA decomposition and *areseq* outputs the ARES model in an explicit mathematical form.

F. Performance Evaluation

The performance of the built ARES models in predicting overall quality score (OQS) and algorithm matching score (AMS) was done. The ARES models were evaluated using known and benchmarked metrics such as model execution time, mean square error (MSE) and coefficient of determination (R2). Other derived metrics like root mean square error (RMSE) and relative root mean square (RRMSE) are also used. The model performance was compared on the training, testing and target datasets respectively and the result was compared with that of [30].

IV. RESULTS AND DISCUSSION

A. FVIQA Implementation

FVIQA was successfully implemented and tested on the SCface database. Its result was found to be consistent with those obtained by [12] [31] [32] [24].

B. Results of Image Quality Scores and Data Pre-Processing

Four score normalization techniques were evaluated in order to determine the most effective and robust method. Table I shows the descriptive statistics of the various normalized data for the faceness measure while Table II shows the result of the correlation analysis between the un-normalized score and the normalized scores for the faceness quality measure. This result is representative of the other quality measures whose data were also normalized. Results on Table II shows that decimal scaling normalization scores had the lowest correlation coefficient R = 0.979 with the original scores. Although the value of R suggests a strong positive linear correlation since it is very close to 1. Fig. 1 shows scatter plots with the correlation between the raw data of faceness measure and the scores of different normalization techniques. Observations on Table I and Fig. 1(c) shows that decimal scaling scores had a much better level of agreement with the original data than the other technique and its scores had better linear spread between the range zero and one. Therefore, scores from decimal scaling normalization technique was adopted as the post-processed quality scores.

Table III presents the statistical parameters of the normalized image quality data for model building while Table IV and V shows the correlation result of overall quality scores (OQS) with individual image quality scores and correlation result of algorithm matching scores (AMS) with individual image quality scores respectively. It was shown on Table IV that pose image quality (QP) had the highest correlation coefficient of R= 0.936 with OQS while on Table V similarity quality (QS) had the highest correlation coefficient of R=0.855 with AMS. The luminance quality (QL) and contrast quality (QC) had the least correlation coefficient for OQS and AMS respectively.

TABLE I: DESCRIPTIVE STATISTICS OF THE RESULT OF EVALUATING THE NORMALIZATION TECHNIQUES

	N	Min	Max	Mean	Std. Deviation
DBE_Un_ Normalized	2936	4.0	57.0	22.2230	10.8457
DBE_Min_Max	2936	0.1527	0.9562	0.3441	0.2048
DBE_Tanh_Est	2936	0.4916	0.5160	0.4999	0.0050
DBE_Z_Score	2936	-1.6802	3.2062	0.0010	1.0006
DBE_Decimal_ Scale	2936	0.0083	1.0	0.4818	0.2345

TABLE II: CORRELATION ANALYSIS OF THE NORMALIZATION TECHNIQUES

		DBE_ Min_Max	DBE_ Tanh_Est	DBE_Dec _Scale	DBE_Z _Score
DBE_Un_ Normalized	Pearson Correl	1.000**	1.000**	0.979**	1.000**
	Sig.	0.000	0.000	0.000	0.000
	N	2936	2936	2936	2936

** Correlation is significant at the 0.01 level (2-tailed)

TABLE III. STATISTICAL PARAMETER OF THE NORMALIZED DATA FOR MODEL BUILDING

Variable	Mean	Std. Dev.	Skewness	Kurtosis
Q'_P	0.5219	0.1883	0.199	-1.503
Q'_L	0.9786	0.0323	-2.819	10.504
Q'_C	0.9555	0.0658	-2.146	4.956
Q'_S	0.0726	0.1999	4.417	17.574
AMS	0.0836	0.2295	3.283	9.702
OQS	0.6021	0.1075	1.347	2.175

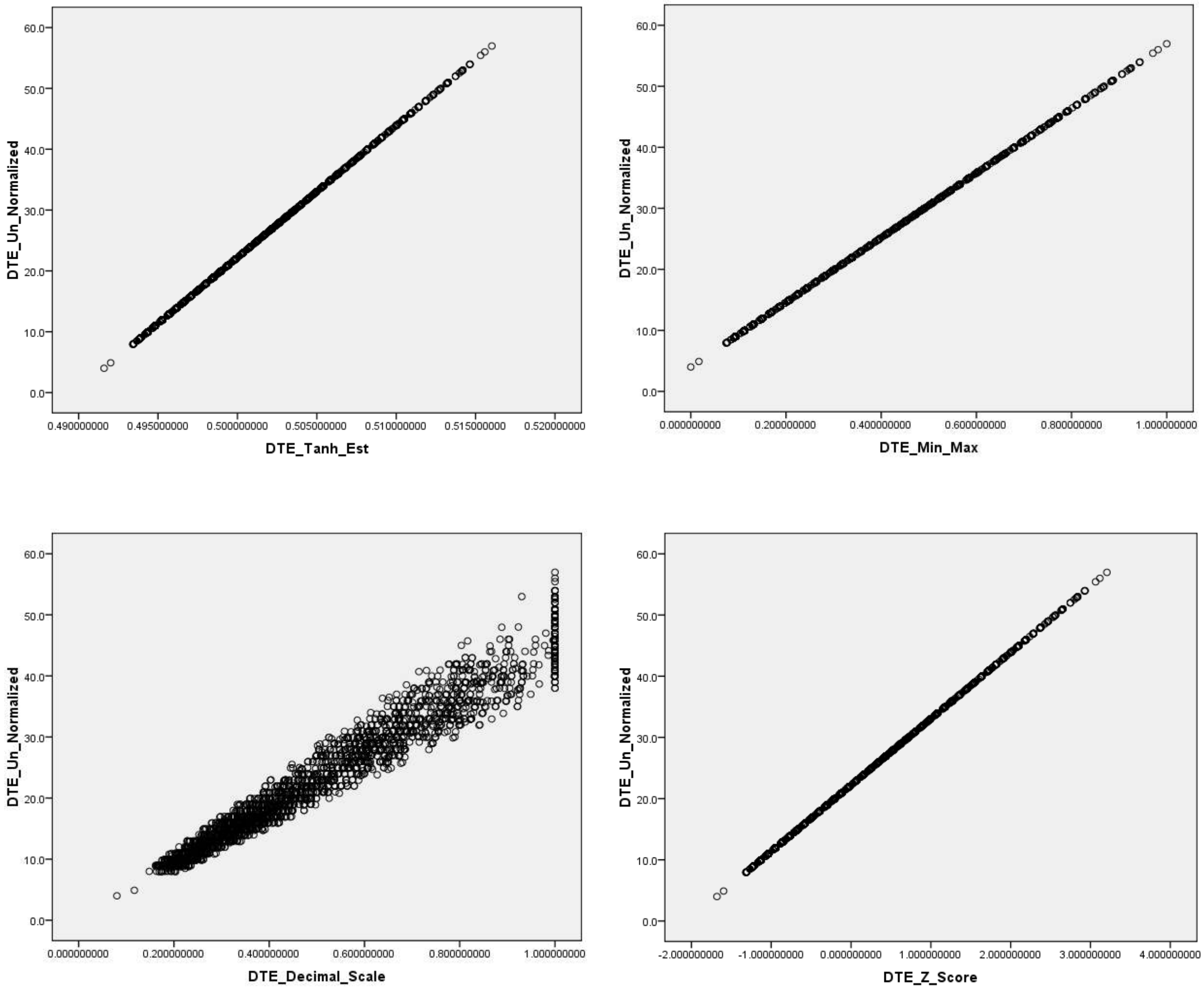

Fig.1. Scatter plots showing the correlation between the raw data of faceness measure and the scores of different normalization techniques

TABLE IV. CORRELATION OF OVERALL QUALITY SCORES (OQS) WITH INDIVIDUAL IMAGE QUALITY SCORES

		QP	QF	QL	QC	QS
OQS	Pearson Correlation	0.936**	0.840**	0.266**	0.262**	0.670**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000
	N	2936	2936	2936	2936	2936

** Correlation is significant at the 0.01 level (2-tailed)

TABLE V. CORRELATION OF ALGORITHM MATCHING SCORES (AMS) WITH INDIVIDUAL IMAGE QUALITY SCORES

		QP	QF	QL	QC	QS
AMS	Pearson Correlation	0.599**	0.379**	0.168**	0.048**	0.855**
	Sig. (2-tailed)	0.000	0.000	0.000	0.009	0.000
	N	2936	2936	2936	2936	2936

** Correlation is significant at the 0.01 level (2-tailed)

C. Adaptive Regression Splines (ARES) Model Building

Once the decimal scaling normalized data was adopted and prepared for the model building, the data was imported into the MATLAB/ARESLab environment from Microsoft office excel. A structure of model configuration parameters was created before calling the function *aresbuild*. These parameters as shown on Table VI were used to build the models OQS and AMS. The data structure of the built ARES models OQS and AMS on Table VII shows that model AMS had more basis functions, longer execution time, higher degree of interactions, higher MSE and GCV values. ARES model OQS had very low MSE of 2.9095E-22 and GCV of 2.9846E-22; this suggests a very good model.

The model ANOVA decomposition on Table VIII for model OQS shows that variable 3 (luminance quality measure) had the least relative importance to the model. The GCV values in the third column however shows otherwise that the ANOVA functions are making equal contributions to the model while Table IX shows clearly that the ANOVA function seven (7) with variables two and three has the least relative importance to the ARES model AMS. This can be interpreted in a manner similar to a standardized regression coefficient in a linear model that variable two (pose measure) and variable three (luminance measure) has the least contribution to determining algorithm matching score (AMS). This is consistent with the results of Table IV and V.

TABLE VI. ARES CONFIGURATION PARAMETERS FOR BUILDING MODELS OQS AND AMS.

Parameter	Configuration	
	Model OQS	Model AMS
maxFuncs	21	21
C	3	3
Cubic	1	1
CubicFastLevel	2	2
SelfInteractions	1	1
maxInterations	2	2
Threshold	0.0001	0.001
Prune	1	1
useMinSpan	-1	-1
useEndSpan	-1	-1
maxFinalFuncs	Infinite	Infinite

Equations 6, 7, 8 and 9 show the mathematical expressions for the models OQS and AMS with their basis functions and coefficients. This makes the deployments of the built models into other image quality assessment framework or software's possible. Due to space constraints we cannot show the explicit mathematical form of the models in terms of the equation cell arrays for the basis functions and corresponding coefficients.

TABLE VII. DATA STRUCTURE FOR THE BUILT ARES MODELS OQS AND AMS.

Data Structure	Model OQS	Model AMS
No. of basis functions	11	14
Total effective No. of parameters	26.0	33.5
Highest degree of interactions in the final model	1	2
Execution time (s)	8.18	22.21
MSE	2.9095E-22	0.0102
GCV	2.9846E-22	0.0105
endSpan	9	9

TABLE VIII. ANOVA DECOMPOSITION FOR MODEL OQS.

Functions	STD	GCV	#Basis	#Params	Variable (s)
1	0.046	0.012	2	5.0	1
2	0.038	0.012	2	5.0	2
3	0.007	0.012	2	5.0	3
4	0.013	0.012	2	5.0	4
5	0.040	0.012	2	5.0	5

Type: piecewise-cubic GCV: 0

TABLE IX. ANOVA DECOMPOSITION FOR MODEL AMS.

Functions	STD	GCV	#Basis	#Params	Variable (s)
1	0.261	0.094	1	2.5	1
2	0.058	0.017	2	5.0	2
3	0.041	0.012	2	5.0	1 2
4	0.056	0.014	3	7.5	1 3
5	0.045	0.013	2	5.0	1 4
6	0.038	0.012	1	2.5	1 5
7	0.011	0.011	2	5.0	2 3

Type: piecewise-cubic GCV: 0.011

D. Model OQS and Model AMS Performance Evaluation

Table 10 shows the result of a 5-fold cross validation for ARES model OQS and AMS respectively. The results shows that model OQS has an average mean square error (avgMSE) of 3.0226E-22 as against model AMS's 0.0115 and an average coefficient of determination (R2) of 0.9996 as against 0.8704 for model AMS. This implies that lower avgMSE gives a better model and the closer the value of R2 to one (1) the better the model. Hence, ARES model OQS had a very strong coefficient of determination with lower error indicators than ARES model AMS. This further suggest that overall quality score (OQS) can be predicted 99.96% accurately by the model OQS given the five quality measures proposed in FaceIVQA while it is only 87.04% for algorithm matching score and model AMS.

Model validation was carried out on the target dataset obtained from the black face surveillance camera (BFSC) database to further validate the performance of the built ARES models. The BFSC database was collected specifically for this research with a difficult experimental setup for the algorithms and a much lower camera quality than SCface. This is to ensure that the data obtained will be a true test of performance for the built models. Results of the test validation for models OQS and AMS on Table 11 shows that model OQS still performed excellently with a MSE of 31.7148E-22, RMSE of 1.3095E-11, RRMSE of 1.2039E-10 and finally coefficient of determination R2 = 0.9981 while model AMS had lower performance score with a value R2 = 0.8473. Comparisons between Tables 10 and 11 shows that despite the test validation on a dataset that is completely different from the training and testing datasets, model OQS still had lower error indicators and a R2 value that is slightly less than 1 while model AMS performance depreciated slightly with higher error scores and a slightly lower R2 of 0.8473 as against 0.8704 on the test dataset.

TABLE X. K-FOLD CROSS VALIDATION FOR MODELS OQS AND AMS.

Performance metric	Model OQS	Model AMS
avgMSE	$3.0226*10^{-22}$	0.0115
avgRMSE	$1.6974*10^{-11}$	0.1072
avgRRMSE	$1.5824*10^{-10}$	0.4677
avgR2	0.9996	0.8704

TABLE XI. TEST VALIDATION FOR MODELS OQS AND AMS ON THE TARGET DATASET.

Performance metric	Model OQS	Model AMS
MSE	$31.7148*10^{-22}$	0.0131
RMSE	$1.3095*10^{-11}$	0.1143
RRMSE	$1.2039*10^{-10}$	0.5027
R2	0.9981	0.8473

Further results of predicting OQS and AMS with the built model are shown on Fig. 2(a) - 2(d) for the testing and target datasets respectively. The predicted scores with models OQS shows close agreement with the actual scores on the testing and target datasets with R2 = 0.9989 and 0.9988 respectively. The plot aligns perfectly with the line of best fit therefore implying accuracy of the model and corroborating the value of R2 as shown on Fig. 2(a) and Fig. 2(b). Therefore, the developed model OQS shows excellent predictive ability for the prediction of image overall quality scores (OQS). For model AMS the results on the Fig. 2(c) and Fig. 2(d) shows a more dispersed spread along the line of best fit but the result is quite acceptable since it is well above 0.8704 and 0.8473 for the test and target datasets respectively.

F. Other Results Obtained

The overall image quality scores (OQS) was categorized into five classes as shown on Table 12 and each IVQA number is a prediction of the recognition algorithm's performance and the contribution of the probe image to the overall performance of the biometric facial recognition system. Fig. 3 shows clearly the implication of this categorization on the experimental dataset. If 1,718 and 1,020 images within the "unacceptable" and "poor" category are discarded from the experimental database then 93.3% (2,738) of the images will be removed and only 6.7% (198) will be left to form a new database. This implies that the new database will contain only images of acceptable (55), good (13) or excellent quality (130).

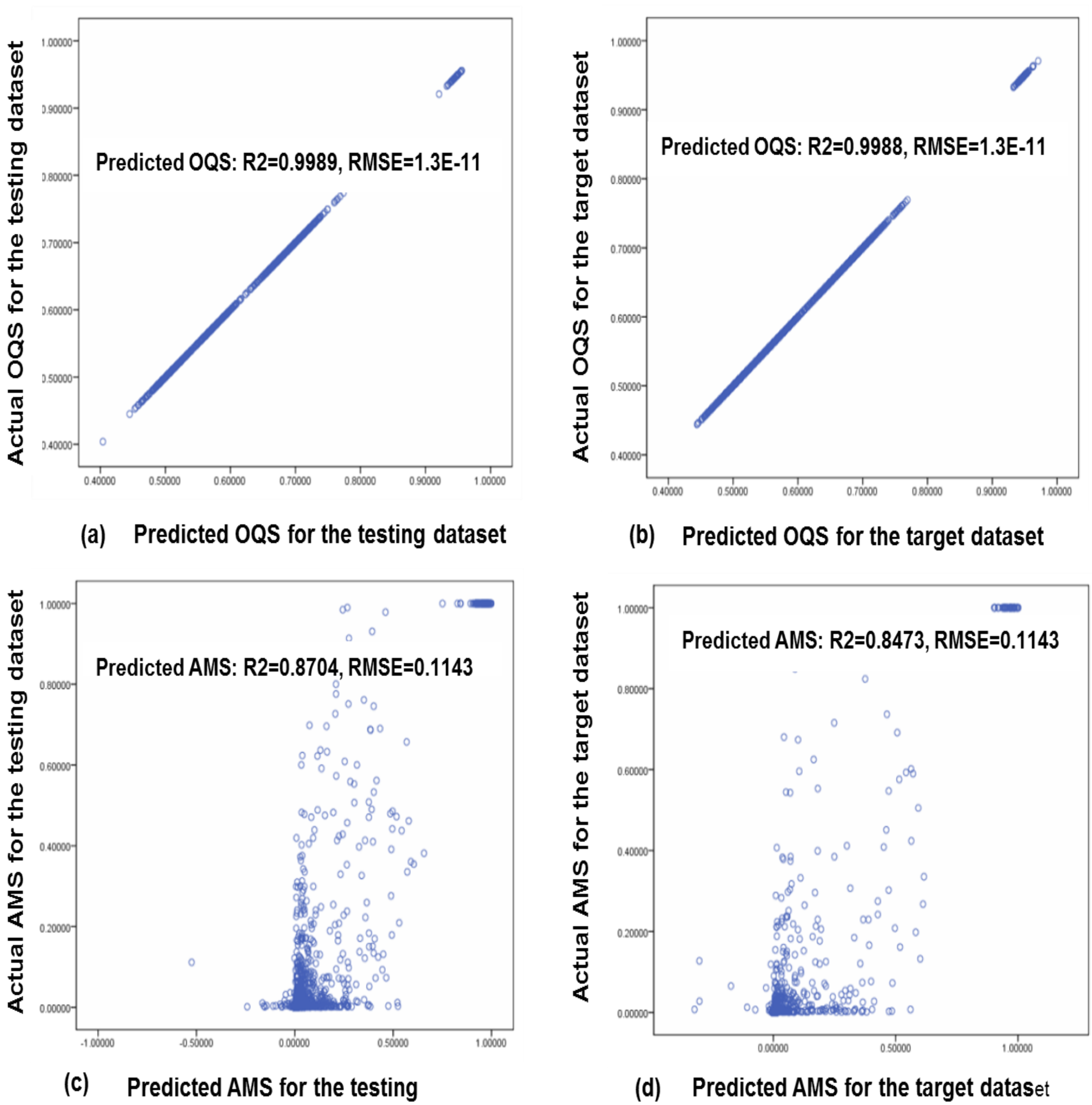

Fig. 2. Performance of overall quality score (OQS) model on the testing dataset

TABLE XII. CATEGORIZATION OF DATABASE PROBE IMAGES

ACROSS QUALITY SCALES.

Overall quality Score range	IVQA number	Description
0.9 - 1.0	5	Excellent
0.80 – 0.89	4	Good
0.60 – 0.79	3	Acceptable
0.40- 0.59	2	Poor
0 – 0.39	1	Unacceptable

TABLE XIII. SUMMARY OF RECOGNITION ALGORITHM'S PERFORMANCE ON THE NEW DATABASE

Algorithm	SR	FTA	TA	FR	FA	TR	MRS
Luxand SDK	198	0	198	0	0	0	0.88
PCA	198	0	198	0	0	0	0.72
LDA	198	0	198	0	0	0	0.67

** Decision threshold = 0.6

Key:
SR = Successful Recognition FTA = Failure to Acquire
TA = True Accept FR = False Reject
FA = False Accept TR = True Reject
MRS = Mean Recognition Score

Hence, the performance of the biometric recognition system will be greatly improved on the new database with 100% accuracy of 198 true accept (TA), zero false reject (FR) and a mean recognition score (MRS) of 0.76 across the three recognition algorithm as shown on Table 13.

Finally, the results presented in this research work are comparable and better to the results reported in [13, 17,21, 30, 33, 34].

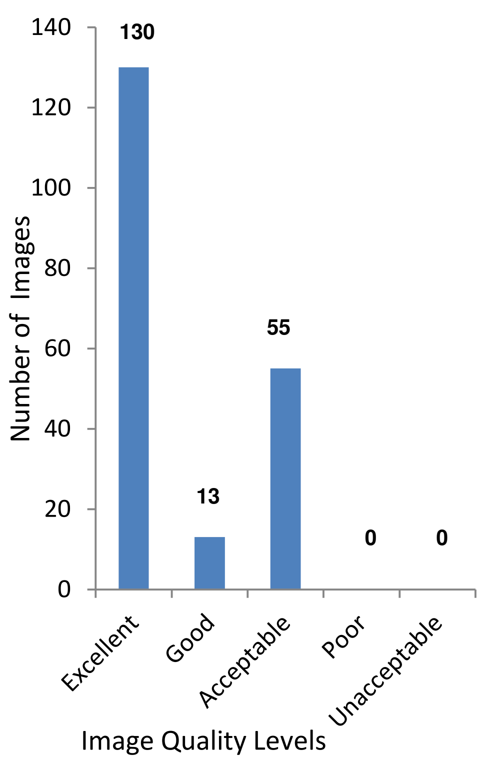

Fig. 3. New database probe images across different quality levels

V. CONCLUSION

Two adaptive regression models for predicting facial image verification and quality assessment scores was built using data from five image quality attributes. Results obtained shows that the individual quality scores are highly correlated with each other and also predictive of the algorithm's matching scores (AMS). This disclosed a correlation between different quality metrics and face recognition performance leading to the possible incorporation of quality measures in reducing the negative effect of poor quality samples. A means of quantifying match performance was developed by combining multiple quality measures into a single overall quality score (OQS). The resulting image verification and quality assessment (IVQA) number can be assigned to images captured for enrollment or recognition for use as input to quality-driven biometric fusion systems.

It was determined that the pose and similarity quality attributes were the most important factors in predicting overall image quality and algorithm matching scores respectively. The models were trained and tested on a 70% to 30% split of the SCface database to avoid systematic errors and ensure that the recognition algorithms as well as the built models were not too sensitive to a mismatch between the training and test conditions. The models were validated on a completely different black face surveillance camera (BFSC) database collected specifically for this research study in order to prove that the recognition algorithms, experimental data and the built models are not in any way bias towards any ethnic or racial group. The overall large size of the probe images used for training and testing was to ensure that random errors which are due to limited number of trials was reduced.

Conclusively, the FVIQA framework adopted has been shown to produce accurate and consistent facial image assessment data in the full-reference verification scenario and the built prediction models in this study using multivariate adaptive regression splines (MARS) technique achieved better performance when compared with some existing works in literature.

We believe this study has contributed the following to the body of knowledge:

a) Development of regression models for predicting facial image overall quality scores and algorithm matching scores using multivariate adaptive regression splines techniques.
b) The explicit model equations can be incorporated and/or employed in other image or pattern recognition problems.
c) A black face surveillance camera (BFSC) database of fifty subjects was collected.

For future research, the following are however recommended:

a) The facial image verification and quality assessment framework should be extended to include other features such as age, color neutrality, exposure, out-of-focus, blur, etc.
b) Automate the data preprocessing stage of the research to allow direct incorporation of the built MARS models into real time quality assessment processes.
c) Preserve all the images in the database while still improving face recognition performance.

ACKNOWLEDGMENTS

This research work is sponsored in part by the Federal Government of Nigeria through the Tertiary Education Trust Fund (TETFUND) academic staff training and development (ASTD) fellowship of the Federal University of Agriculture, Abeokuta, Nigeria.

The authors wish to thank Professors Mislav Grgic, Kresimir Delac and Sonja Grgic for the release of the SCFace-Surveillance camera face database; and the Luxand Corporation for releasing the evaluation version of the Luxand 4.0 SDK library.

REFERENCES

[1] K. Delac and M. Grgic (Eds.), "Face Recognition", I-Tech Education and Publishing, Vienna, July 2007.

[2] S. Z. Li and A. K. Jain (Eds.), "Handbook of Face Recognition". Springer-Verlag, Secaucus, New York, USA, 2005.

[3] K. Delac and M. Grgic, "A Survey of Biometric Recognition Methods", Proc. of the 46th International Symposium Electronics in Marine, ELMAR-2004, Zadar, Croatia, pp. 184-193, 2004.

[4] W. Zhao, R. Chellappa, A. Rosenfeld, P. J Phillips, "Face Recognition: A Literature Survey", ACM Computing Surveys, 35(4):399-458, 2003.

[5] H. Ekenel and B. Sankur, "Feature Selection in the Independent Component Subspace for Face recognition", Pattern Recognition Letters, 25(12):1377-1388, 2004.

[6] N. Poh, J. kittler, S. Marcel, D. Matrouf, J. Bonastre, "Model and Score Adaptation for Biometric Systems: Coping with Device Interoperability and changing Acquisition Conditions", International Conference on Pattern Recognition, IEEE Computer Society, 2010.

[7] U. Park, "Face Recognition: Face in Video, Age Invariance, and Facial Marks", An Unpublished Ph.D Dissertation submitted to the Department of Computer Science, Michigan State University, U.S.A, 2009.

[8] F. Perronnin, "A Probabilistic Model of Face Mapping Applied to Person Recognition", An Unpublished Ph.D Thesis submitted to the Department of Multi-Media Communications, École Polytechnique Federale Lausanne (EPFL), France, 2004.

[9] D. P. D'Amato, "Best practices for taking face photographs and face image quality Metrics", NIST Biometric Quality Workshop, March 2006.

[10] P. Grother and E. Tabassi, "Performance of Biometric Quality Measures," IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(4):531-543, 2006.

[11] F. Hua, P. Johnson, N. Sazonova, P. Lopez-Meyer, S. Schuckers, "Impact of Out-of-Focus Blur on Face Recognition Performance Based on Modular Transfer Function", International Conference on Biometrics, 2012. pp 256-272, 2012.

[12] M. Grgic, K. Delac, and S. Grgic, "SCface - Surveillance Cameras Face Database", Multimed Tools Application, 51:863–879, 2011. Available online at http://www.scface.org/

[13] N. Ozay, Y. Tong, F. Wheeler and X. Liu, "Improving face recognition with a quality-based probabilistic framework", in Proc. Computer Vision and Pattern Recognition (CVPR) conference 2009.

[14] W. J. Scheirer, A. Bendale and T. E. Boult, "Predicting biometric facial recognition failure with similarity surfaces and support vector machines", in Proc. of IEEE Computer Society Workshop on Biometrics, 2008.

[15] M. T. Chan, R. Brown and W. Turner, "Incorporating quality metrics in multimodal biometric fusion", in Proc. of IEEE Computer Society Workshop on Biometrics, pp 67-79, 2006.

[16] K. Nandakumar, Y. Chen, A. K. Jain and S. C. Dass, "Quality-based score level fusion in multibiometric systems", In International Conference on Pattern Recognition, 2006.

[17] P. Wang, Q. Ji and J. L. Wayman, "Modeling and predicting face recognition system performance based on analysis of similarity scores", in Proc. of IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29(4):665–670, 2007.

[18] E. Tabassi, "NIST Fingerprint image quality and relation to PIV", NIST Image Group Technical report, 2005.

[19] J. S. Doyle, "Quality Metrics for Biometrics", M.Sc. Thesis, Department of Computer Science and Engineering, University of Notre Dame, USA, 2011.

[20] E. Tabassi, C. L. Wilson and C. I. Watson, "Fingerprint Image Quality", Technical Report NISTIR 7151, NIST, 2004.

[21] P. Liao, H. Lin, P. Zeng, S. Bai, H. Ma, S. Ding, "Facial Image Quality Assessment Based on Support Vector Machines", in Proc. of International Conference on Biomedical Engineering and Biotechnology (iCBEB), pp. 810-813, 2012.

[22] K Seshadrinathan, T. N. Pappas, R. J. Safranek, J. Chen, Z. Wang, H. R. Sheikh, and A. C. Bovik, "Image quality assessment", Essential Guide to Image Processing, Elsevier, 2009.

[23] C. Sasivarnan, A. Jagan, J. Kaur, D. Jyoti, D. S. Rao. "Image Quality Assessment Techniques on Spatial Domain", IJCST, 2(3):177-184, 2011.

[24] E. O. Omidiora, S. O. Olabiyisi, J. A. Ojo, R. A. Ganiyu, and A. Abayomi-Alli, "Enhanced Face Verification and Image Quality Assessment Scheme Using Modified Optical Flow Technique", in Proc. WCECS'14 International Conference on Signal Processing and Imaging Engineering (ICSPIE'14), San Francisco, USA, 22nd-24th October, 2014.

[25] R. Wallace, M. McLaren, C. McCool and S. Marcel, "Inter-session Variability Modelling and Joint Factor Analysis for Face Authentication", in Proc. International Joint Conference on Biometrics (IJCB), 11-13 October, Arlington, Virginia, USA, pp 1-8, 2011.

[26] J. H. Friedman, "Multivariate adaptive regression splines", Annals of Statistics, 19(1):1-141, 1991.

[27] P.J. Nieto, F. S. Lasheras, F. J. Juez, and J. R. Fernández, "Study of cyanotoxins presence from experimental cyanobacteria concentrations using a new data mining methodology based on multivariate adaptive regression splines in Trasona reservoir (Northern Spain)", Journal of Hazard Matter, 195 (2011): 414–421, 2011.

[28] G. Jekabsons, "ARESLab: Adaptive Regression Splines toolbox for Matlab/Octave", 2011. Available online at http://www.cs.rtu.lv/jekabsons/

[29] P. Craven and G. Wahba, "Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation," Numer Math, Vol. 31:317- 403, 1979.

[30] X. H. Chen and C. Z. Li, "Image quality assessment model based on features and applications in face recognition", in Proc. IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), pp. 1-4, 2011.

[31] E. O. Omidiora, S. O. Olabiyisi, J. A. Ojo, A. Abayomi-Alli, F. Izilein, and P. I. Ezomo, "MACE Correlation Filter Algorithm for Face Verification in Surveillance Scenario", IJCSE, USA, 13(1):6-15, 2013.

[32] E. O. Omidiora, S. O. Olabiyisi, J. A. Ojo, A. Abayomi-Alli, O. Abayomi-Alli, K. B. Erameh, "Facial Image Verification and Quality Assessment System – FaceIVQA", International Journal of Electrical and Computer Engineering (IJECE), Malaysia, Vol. 3(6), 863-874, 2013.

[33] A. Bouzerdoum, A. Havstad and A. Beghdadi, "Image quality assessment using a neural network approach", Proceedings of the Fourth IEEE International Symposium on Signal Processing and Information Technology, pp 330-333, 2004.

[34] S. Bharadwaj, H. Bhatt, M. Vatsa, R. Singh and A. Noore, "Quality assessment based denoising to improve face recognition performance", Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp 140-145, 2011.

BIOGRAPHIES

ADEBAYO A. ABAYOMI-ALLI graduated from the Ladoke Akintola University of Technology (LAUTECH), Ogbomoso, Nigeria in 2005 with a B.Tech degree in Computer Engineering and MSc. degree in Computer Science from the University of Ibadan, Nigeria in 2009. He recently won the Tertiary Education Trust Fund (TETFUND) grant for Academic Staff Training and Development (AST&D) for a one year PhD research bench work at the Sound, System and Structures Laboratory, University of Pittsburgh, USA between May, 2014 and April, 2015. Abayomi-Alli is a registered computer engineer and a chartered IT professional. He is presently a lecturer and researcher at the Department of Computer Science, Federal University of Agriculture, Abeokuta, Nigeria. His research interest includes machine learning, image quality assessment, biometrics and information security.

ELIJAH O. OMIDIORA received the B.Eng degree in Computer Engineering, Obafemi Awolowo University, Ile-Ife in 1992, M.Sc. Computer Science, University of Lagos, Lagos in 1998 and Ph.D Computer Science, Ladoke Akintola University of Technology, Ogbomoso, Nigeria in 2006. He is currently a Professor of Computer Engineering at the Department of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria. He has published in reputable journals and learned conferences. Professor Omidiora is a full member of Computer Professional (Registration) Council of Nigeria (CPN) and a registered engineer by COREN. His research interests are in biometric algorithm and application, image processing and microprocessor based system.

STEPHEN O. OLABIYISI received his B. Tech., M. Tech and Ph.D degrees in Mathematics from Ladoke Akintola University of Technology, Ogbomoso, Nigeria, in 1999, 2002 and 2006 respectively. He also received M.Sc. degree in Computer Science from University of Ibadan, Ibadan, Nigeria in 2003. He is currently an Associate Professor in the Department of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria. He has published in reputable journals and learned conferences. Dr Olabiyisi is a full member of Computer Professional (Registration) Council of Nigeria (CPN). His research interests are in computational mathematics, theoretical computer science, information systems and performance modeling and simulation.

JOHN A. OJO received a B.Tech degree in Electronic and Electrical Engineering from Ladoke Akintola University of Technology (LAUTECH), Ogbomoso, Nigeria in 1998, M.Sc. in the same field from University of Lagos, Nigeria in 2003 and Ph.D from LAUTECH in 2011. He is presently a senior Lecturer in the Department of Electronic and Electrical Engineering in the same institution; he is a member of the Nigerian Society of Engineers (NSE) and a registered engineer by COREN. His research interest includes biometrics, security and surveillance systems, e-health and mobile-health biomedical engineering and image pre-processing techniques.