Informatics in Medicine Unlocked 14 (2019) 43–58

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Implementation of a TCM-based computational health informatics diagnostic tool for Sub-Saharan African students

Olugbenga Oluwagbemi[a,b,*], Abdulwahab Jatto[b,c]

[a] Department of Mathematical Sciences, Private Bag X1, 7602 Matieland, Stellenbosch University, South Africa
[b] Department of Computer Science, Faculty of Science, PMB 1154, Federal University Lokoja, Nigeria
[c] Department of Computer Science, Lens Polytechnic, Km 2 Irra Road, Offa, Kwara State, Nigeria

ARTICLE INFO

Keywords:
Computational health informatics
Traditional Chinese medicine
Diagnostics
Non-communicable diseases
Infectious diseases
Facial image recognition
Informatics

ABSTRACT

Health status checkup is a crucial step towards early detection of diseases. Health status diagnosis, in university health centers, within the sub-Saharan African region, can be cumbersome and time consuming. In many cases, facilities for health checkup are not available. Traditional Chinese Medicine (TCM) is a promising approach, when integrated with *in-silico* methods. This study was conducted to implement a TCM-based computational health informatics diagnostic tool. The tool was applied to diagnose African students. This study was also conducted to stimulate further research into *in-silico* TCM diagnostics. Besides developing a reliable biometric verification system, to ascertain the real identities of patients brought to university health centers, it is assistive to create a platform that provides automated and complementary support for preliminary health diagnostic activities. It also mitigates stress, by helping to efficiently decipher and provide quick objective opinion from the perspective of a computerized decision support system. The diagnostic module of the computational health informatics diagnostic tool adopts knowledge from a TCM facial color diagnosis.

A comprehensive literature search was conducted for relevant full-text research papers. Only research publications written in English language were reviewed. The present work was compared qualitatively and quantitatively with the existing works noted in the literature. Facial detection and matching algorithms were implemented for the TCM-based computational health informatics diagnostic tool by using Java programming language. Facial image acquisition processes were conducted. Captured facial images of African students were preprocessed. Facial feature extraction was performed by implementing feature extraction algorithms. An algorithm for the extraction of color information and measurement was also implemented. Knowledge of machine learning was applied to extract and collate facial features, and to machine learn from them. Facial classification and recognition algorithms were implemented. Finally, the results from the computational health informatics diagnostic tool were evaluated, by conducting a performance evaluation and validation.

This study provides qualitative and quantitative information on facial recognition, facial color information measurement, as well as prediction of health status, for some sub-Saharan African University students. Performance evaluation was shown using confusion matrix and ROC curves. Statistical analysis of the experimental results was presented. The parameters in each diagnostic illustration were shown with valid range. In order to justify the effectiveness of the computational tool, further explanations were provided from relevant methodology guides on the evaluation of diagnostic tests.

The computational health informatics diagnostic tool will complement the diagnostic efforts in university health centers of sub-Saharan African universities. It will also be useful for personal health diagnosis of interested individuals. The tool will also be viable for educating health professionals. TCM will be of immense benefit to developing countries by positively contributing towards diagnosing different non-communicable diseases and some infectious diseases in such countries.

1. Introduction

Health screening and diagnosis are important aspects of medical healthcare. Scientific and technological advancements have led to the development of health screening and diagnosis software. Teo and colleagues [1] highlighted that health screening forms a significant

* Corresponding author.
E-mail address: olugbenga.oluwagbemi@fulbrightmail.org (O. Oluwagbemi).

https://doi.org/10.1016/j.imu.2018.12.002
Received 23 December 2017; Received in revised form 4 December 2018; Accepted 6 December 2018
Available online 15 December 2018

component in disease prevention. Consistent health diagnosis helps to keep track of disease status and associated risk factors. However, health screening is currently not a high priority in many areas of the world, and this has become problematic over time [2–5]. Other problems associated with not participating in health diagnosis include – ignorance, fear of the results of diagnosis, nonchalance towards exercise regimens, and not finding time for getting involved [6,7]. Lack of access to quality health diagnostic facilities, and cost of the screening procedure, amongst others, constitute a substantial barrier toward regular health status diagnosis [8]. It is important to motivate people towards health status diagnosis. This will help prevent a poor state of health, promote early diagnosis of possible diseases, and help prevent untimely deaths.

In some African countries, the process of diagnosing disease can be expensive and cumbersome. This causes delay in the commencement of the requisite medical treatments. There is also a high mortality rate resulting from some chronic diseases. According to the World Health Organization [9], non-communicable diseases (NCDs) have been identified as a leading cause of death globally. Furthermore, diseases such as cancer, diabetes, and cardiovascular heart disease pose great danger to human lives. Mortality rate can be drastically reduced and many deaths averted, if proper, preliminary diagnoses are conducted, before emergency diagnosis and treatment of patients is needed. Thirdly, trained medical doctors cannot diagnose any non-communicable or infectious disease by merely looking observing the patients.

Traditional Chinese Medicine disease diagnosis holds great promise when implemented. This diagnostic approach can be adopted by using computational techniques and methods, which can subsequently be applied in health centers of higher institutions within sub-Saharan Africa. There is a dearth of TCM-related diagnostic research, specifically being applied to the African population. Presently, we are unsure if any TCM works have been specifically applied to disease diagnosis of sub-Saharan African inhabitants. The TCM-based computational diagnostics could be beneficial to indigenous people in developing sub-Saharan African countries.

The aim of this research is to develop and implement a TCM-based computational health informatics diagnostic software. The software will be useful for the diagnosis of African students in sub-Saharan African higher institutions. The objectives of this research are as follows: (i) To implement the methods of a TCM-based approach for computational health diagnostics (ii) To implement a system that can drastically reduce and mitigate the stress encountered by medical personnel (iii) To complement the efforts of medical personnel in the diagnostic processes of non-communicable diseases.

2. Literature review

A comprehensive literature survey was performed for relevant research articles. Many scientific works have previously been conducted on the application of TCM to disease diagnosis. Li and colleagues [10] proposed a Computer Aided Disease Diagnosis System (CADDS) based on TCM, to acquire and analyze facial images for possibly diagnosing disease. Their system collects facial complexion images for the purpose of quantitative-based analysis. Their system consists of a facial image acquisition chamber, a digital camera, and an LED light, used for obtaining accurate facial images of the subjects. The lighting condition is used to overcome the unstable natural lighting in an open environment. Supporting material 1 provides detailed information about each of these literature investigations.

In the study conducted by Li and colleagues, ([11] a facial gloss classification model was developed, based on the knowledge of TCM. They applied a series of feature extraction algorithms on the face gloss (see support material 1). The feature extraction methods were able to extract useful face gloss information. They designed a classification model that produced an automated method for gloss diagnosis. Some of the research gaps in their work include: (i) using feature selection methods to improve the rate of correctness of their proposed system. (ii) using machine-learning methods to improve the rate of correctness of their proposed system.

In another study, Li and colleagues [12] applied TCM-based diagnostics to lip diagnosis of individuals. The observation approach of TCM was adopted in the study. In their experimental results, the best classification accuracy was achieved using Support Vector Machine (SVM).

Observation and inspection are some of the most important procedures in diagnosis [26,27].

Li and colleagues [13] adopted the theory of TCM for clinical diagnosis. One of the most essential procedures in TCM clinical diagnosis, known as '*Wangzhen*, observation or inspection diagnosis' is a facial complexion diagnosis [26,27]. The TCM diagnostics method helps to inspect facial complexion changes, pathological changes and the physiological functions of the human body. Actually, the TCM methodology evaluates areas using five color codes (blue, red, yellow, white, and black). These colors correspond to the liver, lungs, spleen, heart, and kidney respectively. It is believed in TCM that the colors reflect the health conditions of these internal organs. Li and colleagues [13] introduced an acquisition environment for facial recognition and diagnosis based on TCM. Information about facial complexion of subjects was collated. This was done under different artificial light sources. Face color values were extracted. An automated facial complexion-based diagnosis system was built to support quantitative analysis.

Using TCM theory, Zhao and colleagues [14], proposed a new feature representation for the recognition of facial complexion. The results of their study revealed the significance of luminance-level, chromaticity-level and spatiality-level in facial color classification. In addition, they were able to justify that the dominant facial color was more reliably extracted by their two-level clustering method. The results of their research revealed that they achieved a better classification performance for solving facial color problems. Zhang and colleagues [15] explored the relationship between TCM and color images of facial features. The results of their study revealed that LBP textural feature achieved a higher accuracy than the RGB feature. Luo and colleagues [16] conducted a TCM related research. They adopted the pulse palpation approach in their study. Their study investigated how to recognize normal versus hypertensive pulse mappings.

Yang and colleagues [17] conducted a TCM-based study on the extraction of cheek regions within a facial diagnosis process. They selected the cheek region for facial complexion analysis. Liu and colleagues [18], adopted a multi-label learning approach on Coronary Heart Disease (CHD) diagnosis. Three senior TCM doctors performed CHD diagnosis on each of 555 patients. Diagnosis criteria in Western Medicine and TCM were adopted. The symptoms in eight dimensions were: cold or warm, sweating, head, body, chest and abdomen, urine and stool, appetite, sleeping, mood, and gynecology. There were 15 syndromes in differentiation diagnosis. They constructed symptoms-syndromes relationship models using multi-label k-nearest neighbor (MLkNN) and classical k-nearest neighbor (kNN) algorithms. The results from both algorithms were compared. ML-kNN produced better results than RankSVM, BPMLL and kNN. Mist and colleagues [19] conducted a study on the effects of training and questionnaire-based diagnosis on inter-rater reliability of 10 TCM practitioners. The practitioners were those evaluating patients with temporomandibular joint disorder (TMJD).

Lo and colleagues [20], applied logistical regression and the Mann-Whitney test to data collected from early Breast Cancer (BC) patients. They applied the TCM observation approach towards tongue diagnosis. Accuracies of 80%, 80% and 90% were obtained for non-breast cancer individuals. Accuracies of 60%, 60% and 50% was obtained for early BC patients. The results obtained in the study conducted by Kang and colleagues [21] showed that higher syndrome classification performance was based on a combination of TCM and MM modern clinical indices. Xue and colleagues [22], adopted the stratification, treatment, observation, assessment, and statistical analysis approaches on the data of elderly patients with advanced NSCLC. They conducted a comprehensive assessment and traditional Chinese medicine intervention benefit on the patients. An integrated facial feature was proposed for the diagnosis of the Chronic Fatigue Syndrome (CFS) by Chen and colleagues [23]. This hybrid feature was based on the observations made by TCM doctors. An observation TCM approach was adopted in the study. Li and colleagues [24] applied a new multi-label learning model to process the clinical data of hypertensive patients. The clinical data in their study was collated through inspection, inquiry, and palpation.

Watsuji and colleagues [25], developed a fuzzy diagnostic system for tongue inspection. The diagnostic system was able to diagnose several syndromes. The diagnostic system was found to be useful for tongue inspection. Moura and colleagues [26], carried out an evaluation, by relating PIA with hemodynamics and PWA. They conducted their research on patients with hypertension. A certified TCM practitioner conducted the analysis of patients. They adopted the TCM approach of observation, pulse feeling and palpation. Jiang and Liang [27], conducted a study on olfactory diagnosis of human subjects. They adopted the TCM approach for olfaction, to establish olfactory diagnosis for the patients involved. Xu and colleagues [28] obtained four types of diagnostic data from 835 CHD patients. TCM approaches were adopted through inquiry or interrogation, pulse feeling, palpation, auscultation, olfaction, and observation. A multi-label learning algorithm was used for syndrome classifications. In the research conducted by Wu and colleagues [29], healthy humans were examined by applying a pulse wave.

Zheng and colleagues [30] conducted a study on the application of questionnaire administration and TCM-based diagnostic approaches to identify better treatments for stressed patients. The study conducted by Zhang and colleagues [31] focused on the diagnosis of diabetes, by adopting an integrative knowledge of SVM and images of the tongue. A TCM observation approach was adopted. GA-SVM was found to be a more efficient and a better classification model for tongue manifestation than kNN, Naive Bayes, and the Backpropagation Neural Network (BP-NN). Tian and colleagues [32] investigated how the integration of TCM and Western Medicine can assist in improving the diagnosis and treatment of knee osteoarthritis. Their results revealed TCM modes that could help improve diagnosis and treatment of knee osteoarthritis. In the study conducted by Jiang and colleagues [33], tongue-image diagnosis and analysis were conducted. Feature extraction from tongue images was performed. Eighteen digital features were extracted and analyzed using Principal Component Analysis (PCA). DNA samples were extracted and analyzed. Sequencing data and statistical analysis were performed. Dissimilarity measures among the samples was measured using the Jaccard, Bary-Curtis, unweighted Unifrac and weighted Unifrac distance measures.

Wang and Cheng [34], formulated a pulse diagnosis model, based on Bayesian Networks (BNs). They adopted TCM approaches of pulse diagnosis and palpitation. The predictive capacity of their system had an accuracy of 84%. Chiu and colleagues [35], conducted experiments on the digitalization of speech signals for healthy and deficient individuals. They performed some data analysis. They adopted four novel acoustic parameters. The average number of zero-crossings, the variations in local peaks and valleys parameters, outperformed other parameters. O'Brien and colleagues [36], assessed the treatment efficacy of Chinese medicine on an Australian population with hypercholesterolemia and other cardiovascular risk factors. They conducted an evaluation on the reliability of three of the TCM diagnostic approaches. Results showed that certain TCM features were repeatable, while other features were unreliable. Hua and colleagues [37] conducted an assessment of Chinese medicine diagnostic variables in the study of patients with knee osteoarthritis. Two TCM doctors and forty patients were involved in the study. Data collation, assessment of CM diagnostic variables, and statistical analysis of the levels of agreement among the variables, characterized the method of the study. Some variables had a higher agreement, while others had a lower agreement. Ferreira [38,39] conducted a study on the diagnostic accuracy of patterns inherent in collated datasets. A stochastic simulation study was also conducted on the similarity of patterns inherent in Chinese Medicine collated datasets. They adopted TCM methods of inquiry, inspection, auscultation/olfaction, and palpation.

Jeon and colleagues [40], conducted a quantitative study by analyzing the parameters associated with pulse diagnosis of 20 healthy subjects at three different positions (Chon, Gwan and Cheok). The analysis was done by studying the behavior of the parameters. They adopted the TCM approaches of pulse diagnosis, pulse feeling, auscultation and palpation in the study. Hui and colleagues [41] applied five machine learning algorithms to analyze tongue datasets. They adopted the TCM approach of inspection. Their results showed that the Support Vector Machine SMO algorithm had the best performance in analyzing the tongue dataset. SMO had the highest accuracy, with a cross-validation of 93.45%, 96.03% leave-one-out, and 90.38% of percentage split. The Support Vector Machine-based Sequential minimal optimization (SMO) algorithm, also achieved the best average Area Under the Receiver Operating Characteristics Curve (AUC) performance. ROC can be used to measure the performance of algorithms. Yuen and colleagues [42] engaged in a TCM study by adopting a computer vision technique for tongue diagnosis. They also used Gabor Wavelet Opponent Color Features (GWOCF) to determine the tongue texture. They applied color information to pre-classify tongue texture.

Bakshi and Pal [43], engaged in a TCM tongue diagnosis study that involved inspection and observation. Facial and tongue images were captured before treatment and analysis. Other biological samples (blood, stool and urine) were collected for examination. Images were later analyzed for correlation. Hu and colleagues [44], applied TCM pulse diagnosis to study pulse differences in different human subjects. The Bi-Sensing Pulse Diagnosis Instrument (BSPDI) was applied in the diagnosis process. A pulse taking procedure was also done using the Three Positions Nine Indicators (TPNI) method. Clinical analysis/diagnosis was conducted. Sampling of wrist signals, computing the wrist pulse signals, signal analysis, and the construction of 3D pulse mapping by surface fitting equations, were conducted. Anastasi and colleagues [45], conducted a TCM tongue diagnosis study on HIV patients. Tongues of patients were examined. Patients were asked to fill in initial assessment forms; questionnaires were also administered to patients. The data collated were analyzed statistically. Zhao and colleagues [46], adopted TCM methods to identify and diagnose major syndromes in patients living with chronic hepatitis B (CHB).

A detailed, but tabulated, summarized version of the systematic study of TCM related works can be found in supporting material 1. There is also a brief comparative analysis of our work with other existing TCM-related works in the table in supporting material 1.

Other health-related diagnostics applications include Ebinformatics developed by Oluwagbemi and colleagues [47], which focused on diagnosing the Ebola Virus Disease (EVD) based on symptoms keyed into the system by patients or medical personnel. The diagnosis module of Ebinformatics was developed by applying the Fuzzy inference engine, Fuzzy rules, Fuzzy sets, and defuzzification mechanisms. In another positive development, Oluwagbemi and colleagues [48], developed a mobile application that provides comprehensive knowledge to users about rare and common hereditary diseases. They developed the diagnostic module in the mobile application by using the Logical Disjunction Rule-based Algorithm (LDRA). Oluwagbemi and Oladunni [49] developed a web-based diagnostic and recommender system, based on a client/server architecture, for some neglected tropical diseases. The knowledge of data mining has been applied to the development of diagnostic systems. Oluwagbemi and colleagues developed a knowledge-based data mining system to diagnose malaria cases in the management of healthcare [50]. Computer-Based Disease diagnostic systems have been developed in the past. An expert system for malaria environmental diagnosis has been previously developed by Oluwagbemi and colleagues [51]. They integrated the knowledge of Java programming, Java Expert Systems Shell (JESS) and SQL server to implement the diagnostic system. A diagnosis module was developed in *Malavefes* by Oluwagbemi and colleagues to predict malaria intensity in

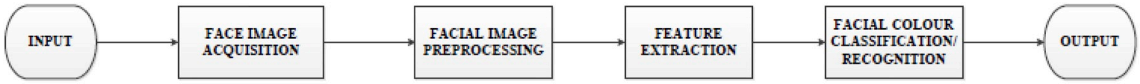
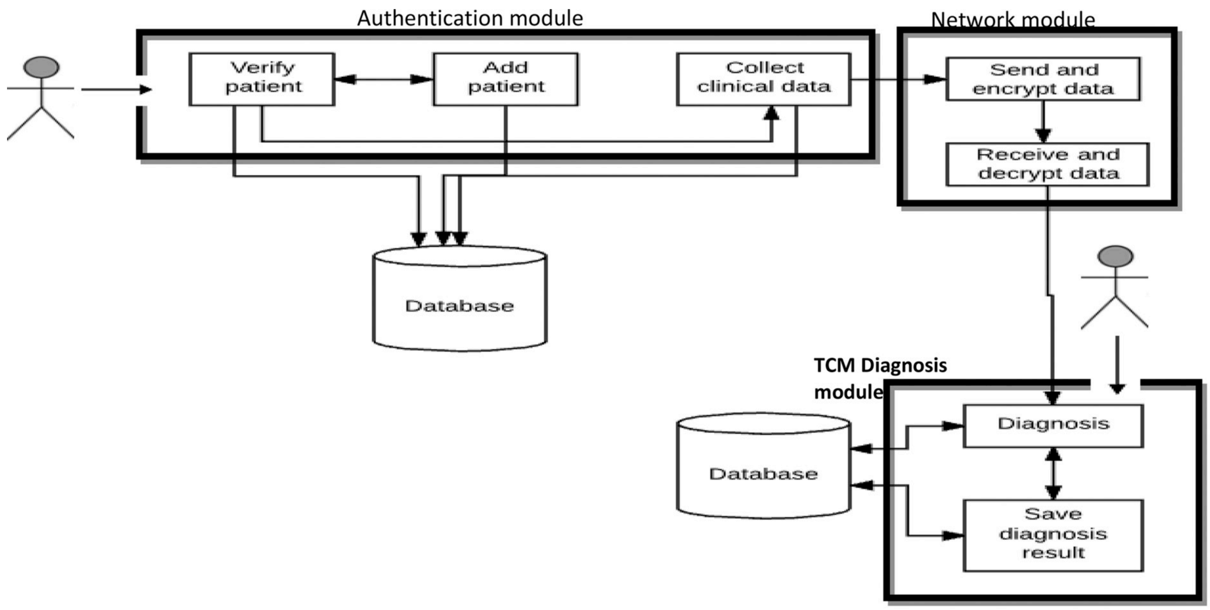


Fig. 1. Facial diagnosis and classification process.



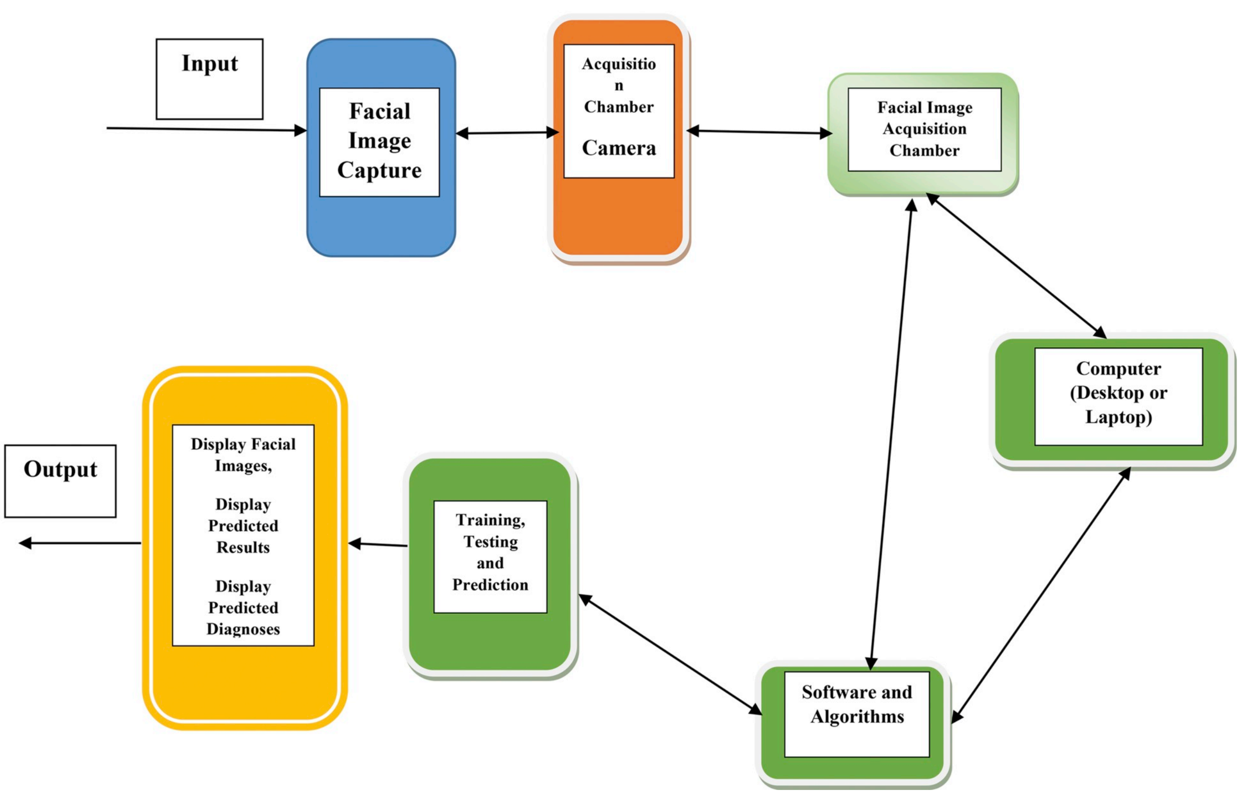


Fig. 2. A schematic/block diagram which shows the interconnection of the facial acquisition chamber and components of the whole system.

patients [52].

3. Methodology of experiments

This methodology for this research was divided into three phases, namely: (i) the verification and documentation phase (ii) the diagnosis phase and (iii) the network phase (See Fig. 2a, in the Supplementary material 2). The GUI of the software, diagnosis, and verification phases were implemented using Java programming language. The knowledge of TCM, facial complexion, and color diagnosis was adopted. This knowledge was implemented by using computational techniques. This required the acquisition of facial images, implementation of the preprocessing phase, extraction of facial features, and finally, facial color classification and recognition.

3.1. Verification and documentation phase

In the TCM-based computational health informatics diagnostic software, this module captures the biometric features of the patients, and helps to verify them by recognizing the facial attributes of the patient. This phase helps the health facility to recognize and verify the identity of the patient.

At the input stage, an image is placed in front of the software's camera as data. This is followed by facial detection of the incoming data. This involves searching for and identifying the face evident in the image. Face preprocessing is the next stage. It involves cleaning the image identified for easy recognition. The next stage is the collection and learning stage. This involves the capturing, saving of many preprocessed faces, and learning to recognize the images. Facial recognition is the next stage. It involves conducting a comparative analysis between the preprocessed facial image and a repository of known faces, to determine and reveal the true identity of the patient.

3.2. Network phase

One of the problems in university health centers in countries within the Sub-Saharan African region, is that some of these health centers still handle medical records on paper files. Manual operations occur in such health institutions through the handling of medical records in paper files. The absence of computer networking facilities in such health centers means that medical records are not networked. In our TCM-based computational health informatics diagnostic software, we have made provisions for the secure transfer of patient data among medical personnel within the health center.

Data collated during preliminary physical examination of patients, by nurses, are sent to the physician. The data is collated for further analysis, via an internet/intranet, within the health center. In order to make this data secure, this phase implements the 128 bit key-size Advanced Encryption Standard (AES) [53–56] and image steganography [57–59]. Such data can be used along with the TCM facial diagnosis data to provide a more comprehensive diagnosis of patients.

3.3. Diagnostic phase

The diagnostic phase involves six stages: the input stage, facial image acquisition stage, facial image preprocessing stage, feature extraction stage, facial color classification/recognition stage, and output stage (Fig. 2). This module performs TCM facial color diagnosis, also known as the TCM facial complexion diagnosis. The facial diagnosis is performed on the subject by computerized inspection. In this work, computerized facial image analysis was conducted, on experimental subjects' facial image in an enclosed facial image acquisition chamber. The figure below depicts the stages in TCM facial complexion diagnosis in our work (see Fig. 1).

3.3.1. Experimental subjects

Two (2) human experimental subjects volunteered for this experiment. The inclusion criteria for students include: age between 17 and 35 years. The gender could be male or female. Smoking status could be cigarette or non-cigarette smokers. Drinking status could be alcohol or non-alcohol drinkers. The subjects gave their consent. These were university students between the ages 20–24 years. One of the participant was a male, while the other was a female. The two students were non-smokers and non-alcohol drinkers. We weren't able to obtain consent of African students that smoke cigarettes or drink alcohol to participate in the experiments. The facial images of these subjects were captured using the TCM-based acquisition chamber. The TCM-based acquisition chamber was locally constructed and connected to a computer. The schematic diagram of the entire system is depicted in Fig. 2.

3.3.2. Input

The input stage involves placing the face of the patient in front of the facial acquisition chamber. The chamber is a box-like compartment that contains a digital image acquisition/capturing device, red electric bulbs, green electric bulbs, blue electric bulbs, yellow electric bulbs, white electric bulbs, and natural light and darkness sections. The chamber also consists of attached electrical wires, as well as switches to control lighting effects. The facial image acquisition chamber also consists of a smaller hole for placing and holding a high resolution USB digital camera. Other components in the chamber include: resistors, batteries, electrical wires, battery terminal connectors, switches, LED (Light-Emitting Diode)/bulbs and PCB (Printed Circuit Board).

The Facial acquisition chamber has an embedded camera and a space for capturing the facial image of a patient. The facial acquisition chamber is connected to a computer (the computer contains the TCM-based computational health informatics diagnostic software).

3.3.3. Facial image acquisition

We designed and constructed a facial image acquisition chamber for this purpose. The facial image acquisition chamber consists of a large opening to encompass the surface area of the face for image acquisition purpose. The internal surroundings of the facial image acquisition chamber were designed to be able to suspend and hold an LED light (of different colors like red, blue, green, yellow, white light), in a perpendicular direction to the camera, in order to acquire the accurate facial image for subjects, with the correct lightning condition.

Facial image samples, with different colors of light, were collected by an enclosed facial image acquisition chamber. The performance of the facial image analysis, of the TCM-based computational health informatics diagnostic software, can be affected by the quality of the facial image acquisition devices. There are certain factors that can affect facial image quality. They are (i) the quality of lightning/illumination condition of the environment, and (ii) the quality of the facial image camera and chamber. Zhao and colleagues [14], developed a framework for facial image acquisition. The framework entailed the projecting of different colors on a subject's face. The colors are as follows: red, green, blue, yellow, normal sunlight, and black. Zhao and colleagues [14] recommended that the TCM facial complexion diagnostic framework could be used by people of other races. We adopted part of this knowledge, from the framework proposed by Zhao and colleagues, in our computational health informatics diagnostic software.

Facial detection with OpenCV (Open Source Computer Vision), Adaboost trained cascade classifier, was implemented (see listing 1: check the facial detection algorithm [71]). Detected facial image samples were collected by an enclosed facial image acquisition chamber. These samples were further preprocessed for color space transformation and clustering. Thirdly, features were extracted from the preprocessed samples to obtain image patterns. Lastly, the patterns from the preprocessed image samples were utilized for facial color classification/recognition.

3.3.4. Facial image preprocessing stage

Samples were preprocessed for color space transformation and clustering. The study of color spaces is significant to facial image recognition ([60]. In order to retain the facial image color and textural information, the Commission Internationale de l'Eclairage (CIE) illuminant D65 was adopted. Here, each facial image color space was captured in the RGB color space, then transformed to CIEXYZ color space, and then transformed to the CIELAB [72]. [see Listing 2 for the color transformation RGB to CIEXYZ to CIELAB]. The reason for image color space transformation is that digital camera images suffer a device dependent color space rendering [61].

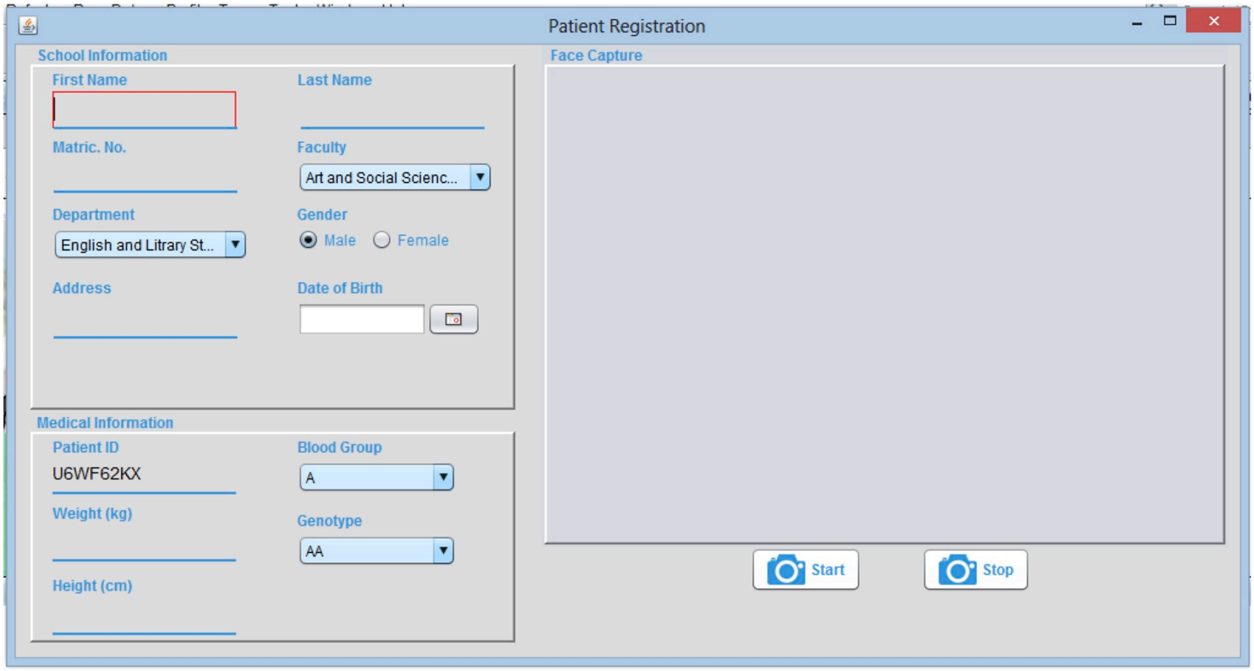


Fig. 3a. Patient registration form.

3.3.5. Feature extraction stage

Feature extraction was conducted on the preprocessed image samples [[62]. In the TCM approach, there exists a common dominant color, equally proportioned, and distributed over the entire face [61]. Indigenous TCM practitioners believe that the dominant condition can reflect health conditions. So during their practices of disease diagnosis, they extract the chromatic dominant color. Our TCM-based computational health informatics diagnostic software was able to extract color, texture, and algebraic features from each facial image sample. The total vector length extracted from each facial image sample was 36 elements. ColorSummarizer [63], a perl framework developed by Martin Krzywinski for image analysis, was used to extract the color feature. ImageJ [68], a java framework for medical image analysis [64–66], was also applied.

In TCM practice, there is usually a common color largely distributed over the entire face. The color is depicted as the dominant color [61]. It is the belief of TCM doctors that the dominant color can, to some extent, reflect the health condition of patients. Chromatic dominant color can also be extracted by TCM doctors during diagnostic processes.

Extraction of color, textural features, and algebraic features were performed on the clustered facial image. In order to extract the color features from the clustered facial image, the system transforms the clustered facial image color space into other sets of color spaces, like HSV (Hue Saturation Value), CIELCH, and CYMK. A vector length of 19 color features was obtained from the clustered facial images. [see the listings 2, 3 and 4, on color feature extraction]. The textural feature extraction techniques implemented were developed by Haralick and colleagues [74]. These include entropy, contrast, correlation, and variance. A vector length of 14 textural features was obtained from the clustered facial image [see the texture features that can be extracted in listing 3]. Extraction of algebraic features from sample images was performed using the SVD [75] [see Listing 4 - the algebraic feature extraction algorithm]. The clustered facial image produced a vector length of 3 algebraic features.

Eventually, a total vector length of 36 was obtained from each sample image. Patterns extracted from each sample image are depicted by these images. These patterns act as inputs into the SVM classifier. The color feature was extracted using ColorSummarizer [63]; a perl framework for image analysis. The textural feature of each image sample was extracted using ImageJ [68]; the algebraic features were extracted using Apache commons-math3-3.2 Java library [75].

3.3.6. Facial color recognition and classification

We utilized the features extracted from the preprocessed image samples for facial color recognition and classification [62]. In this phase, our TCM-based computational health informatics diagnostic software adopted the Support Vector Machine (SVM) library- LIBSVM (a Library for Support Vector Machines originally created by Chih-Chung Chang and Chih-Jen Lin [69]), a machine learning algorithm to classify the facial color of the trained and tested sample images, (See Jatto test.txt, Jatto train.txt, Stella test.txt, Stella train.txt). The mined patterns were classified from the training and testing image samples.

3.3.7. Output stage

The sixth and final stage is the output stage. Here, the final results of facial recognition/classification is outputted and used for patient disease diagnosis. The output of our TCM-based computational health informatics diagnostic software can either be indicated as red, yellow, black, blue, green, or normal (i.e. the healthy color) [See Fig. 3e, f, and 3h - Results section].

3.4. Implementation and experiments

The experiment adopted the LIBSVM (Library for Support Vector Machines) was originally created by Chih-Chung Chang and Chih-Jen Lin [69]. We actually used a more recent version of the LIBSVM [70] Chang and Lin, 2016 version 3.22) which was implemented in the Java programming language. LIBSVM implements the SVM algorithm, which is both an easy and efficient framework for solving medical image classification problems. The system collates 6 training image samples and 1 test image sample; this sums up to 7 sample images per patient (see Jatto test.txt, Jatto train.txt, Stella test.txt, Stella train.txt).

Thus, the implementation tools for these experiments are: Colorsummarizer-0.77 (an image color summarizer tool developed by Martin Krzywinski), Netbeans 8.1 and above, JavaCV 0.01, ImageJ, LIBSVM, MySQL Database management system and sqlite 3.

Listing 1. Facial detection algorithm (Source: Wang, 2014)

Input *an MXN grayscale image I and an L-layer cascade of shifted classifiers trained using Attentional cascade*
Parameter: *a window scale multiplier c*
Output: *P, the set of windows declared positive by the cascade*
Set P $= f[i; i+e-1]_X[j; j+e-1] \subset I : e = J24c^k K; k \in \mathrm{N}$
For *I = 1 to L do*
For *every window in P do*
Remove *the windowed image's mean and compute its standard deviation*
If *the standard deviation is bigger than 1 then*
divide *the image by this standard deviation and*
compute *its features required by the shifted classifier at layer I with Feature Scaling*
If *the cascade's I-th layer predicts negative then*
discard this window from P
end If
else
discard this window from P
end If
end For
end For
Return P

Source: (Wang, 2014) [71].

Listing 2. : Feature extraction algorithm

Feature Extraction Algorithm.

Formulas governing Algorithm converting RGB to CIEXYZ to CIELAB.

Color Features for color space transformation

$RGB \rightarrow CIEXYZ \rightarrow CIELAB$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 \\ 0.2126 \\ 0.0193 \end{bmatrix} \begin{bmatrix} 0.3576 \\ 0.7152 \\ 0.1192 \end{bmatrix} \begin{bmatrix} 0.1805 \\ 0.0722 \\ 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

Then, the transformation from CIEXYZ color space to CIELAB color space is done by using the following formulas:

$$L^* = 116f\left(\frac{Y}{Y_n}\right) - 16$$

$$a^* = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right]$$

$$b^* = \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \tag{2}$$

where

$$f(t) = \begin{cases} t^{\frac{1}{3}} if t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29} otherwise \end{cases} \tag{3}$$

CIEXYZ tristimulus values of the reference white point are X_n, $Y_n andZ_n$. Often, its values are assumed asX = 95.047, Y = 100, and Z = 108, 883 relative to CIE standard illuminant D65.

Source [14]: Zhao et al., 2014

Equations governing the conversion of RGB to CIELAB to CIELCH

$RGB \rightarrow CIEXYZ \rightarrow CIELCH$

The equations highlighted in 1, 2 and 3 depict the conversion of RGB to CIEXYZ to CIELAB.

In order to convert RGB to CIELAB, and then from CIELAB to CIELCH, we have:

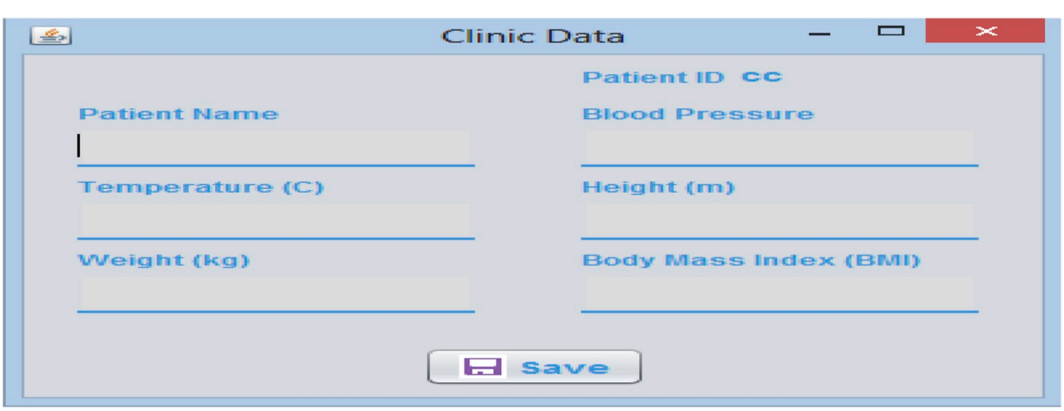


Fig. 3b. Preliminary Examination Clinic Data Form; Click on the Clinical Data icon to collect patients' clinical data.

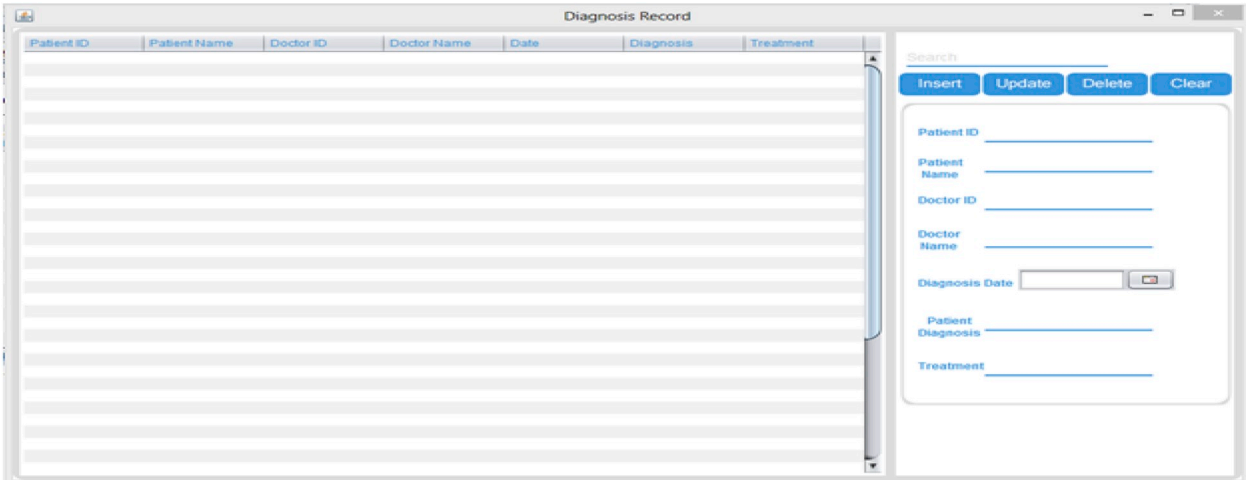


Fig. 3c. Click on the Records icon to search, update, insert, and delete diagnosis records.

L = *L*

C $= \sqrt{a^2 + b^2}$

$$\mathbf{H} = \begin{cases} \arctan\left(\frac{b}{a}\right), \arctan\left(\frac{a}{b}\right) \geq 0 \\ \arctan\left(\frac{b}{a}\right) + 360, \textit{otherwise} \end{cases}$$

RGB → *HSV*

V = *Maximum (R, G, B)*

Delta = *V - Minimum(R, G, B)*

S = *choose (V=0, 0, Delta * 255 / V)*

If *S = 0*

H = 0

else If

Red = V

H0 = (Green - Blue) / Delta

elseif

Green = V H0 = 2 + (Blue-Red) / Delta

else

H0 = 4 + (Red-Green) / Delta

end

If *H0 < 0*

H0 += 6

end

*H = H0 * 255/6*

End

GB → CMYK

C = 255 - Red

M = 255 - Green

Y = 255 – Blue

K = Minimum (C, M, Y)

Source [73]: (Nishad and Chezian, 2013).

Listing 3. : Texture Features Extraction

Textural features that can be extracted from images include:

(1) Angular Second Moment: $f_1 = \sum_i \sum_j \{p(i,j)\}^2$

(2) Contrast: $f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\}$

(3) Correlation: $f_3 = \frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$

Where, μ_x, μ_y, σ_x, and$\sigma_y$are the means and standard deviation of. $p_x$andp_y

(4) Sum of Squares: Variance

(5) Inverse Difference Moment: $f_5 = \sum_i \sum_j \frac{1}{1+(i-j)^2 (p(i,j).}$

(6) Sum Average: $f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i)$

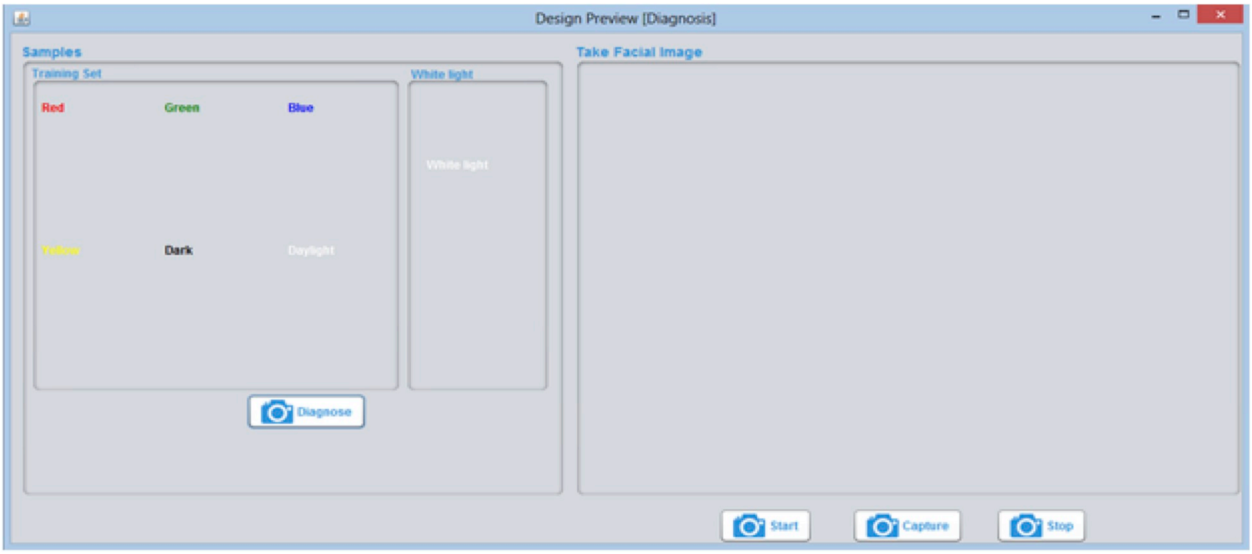


Fig. 3d. Diagnosis Form: click on the diagnosis form to diagnose patients.

(7) Sum Variance: $f_7 = \sum_{i=2}^{2N_g}(i - f_8)^2 p_{x+y}(i)$

(8) Sum Entropy: $f_8 = -\sum_{i=2}^{2N_g} p_{x+y}(i)\log\{p_{x+y}(i)\}$.

(9) Entropy: $f_9 = -\sum_i \sum_j p(i,j)\log(p(i,j))$

(10) Difference Variance: f_{10}=variance of p_{x-y}

(11) Difference Entropy: $f_{11} = -\sum_{i=0}^{N_g-1} p_{x-y}(i)\log\{p_{x-y}(i)\}$.

(12),(13) Information Measures of correlation: $f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}}$

$$f_{13} = (1 - \exp[-2.0])(HXY2 - HXY)^{1/2}$$

$$\mathrm{HXY} = -\sum_i \sum_J p(i,j)\log(p(i,j))$$

Where HX and HY are entropies of $p_x$and p_y, and

$$\mathrm{HXY1} = -\sum_i \sum_j p(i,j)\log\{p_x(i), p_y(j)\}$$

$$\mathrm{HXY2} = -\sum_i \sum_j p_x(i), p_y(j)\log\{p_x(i), p_y(j)\}$$

(14) Maximal Correlation Coefficient:

$$f_{14}(Second\ l\arg est\ eigenvalue\ of\ Q)^{1/2}$$

where. $Q(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}$

Source [74].

3.4.1. *Algebraic feature extraction*

Singular Value Decomposition (SVD) can be used to connote

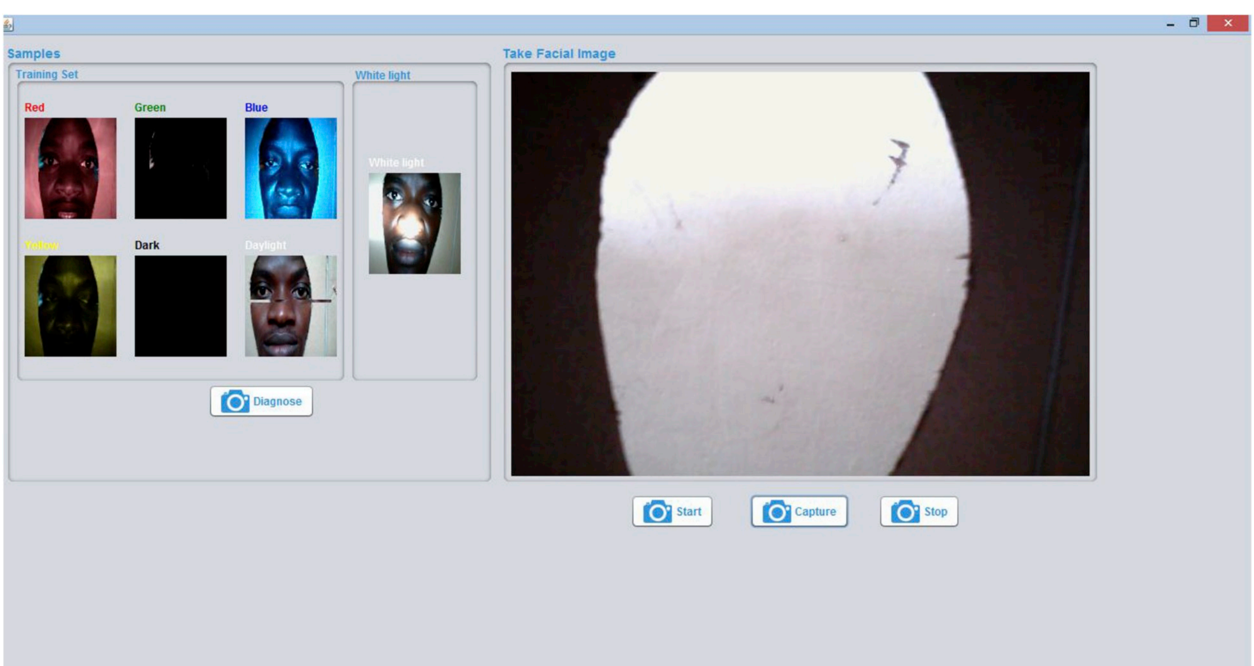


Fig. 3e. Facial image samples captured by the computational health informatics diagnostic software for student 1 in one of the Sub-Saharan universities.

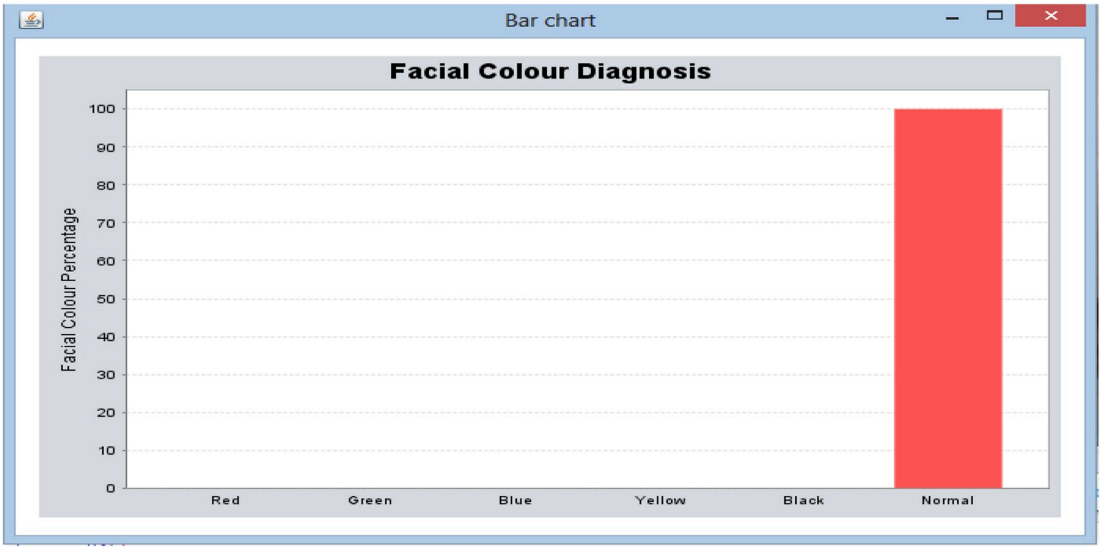


Fig. 3f. Facial Color Diagnosis Classification graph depicting output of the TCM-based computational health informatics diagnostic software for student 1. Here, the red indicator points at the normal segment which indicates that the patient is healthy. In TCM theory, five (5) internal organs; liver, lungs, spleen, heart, and kidney, are mapped to five colors; blue, red, yellow, white and black respectively, which may appear on the face. These colors show the conditions of these organs [13]. In our own results, blue, red, yellow, white, and black represents liver, lungs, spleen, heart and kidney, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

algebraic features of an image, which are inherent and not evident. SVD has been applied in different domains, namely: data compression, pattern recognition (PR) and analysis, and signal processing. Singular Values (SV) from image feature have attributes of both algebraic and geometric invariances [75] (Hong, 1991).

Listing 4. : SVD algorithms

Let *A_ (m×n) be a rectangular matrix and Rank (A)=k*

Two *orthonormal matrix U_(m×m) , V_(n×m) and Σ_(m×n) exist.*

Compute *A=UΣV^T where Σ=diag (λ_1, λ_2,…,λ_k,0,…,0).T denotes transpose.*

SV *features vector is unique to A*

End

Source [75].

4. Results

The result of this research was to produce a TCM-based computational health informatics diagnostic software. The results produced from the computational health informatics diagnostic software can be found in Fig. 3a–h. Table 4 shows the results obtained from the computational health informatics diagnostic tool, when compared with the medical results obtained for the experimental subject (Student 1). Table 5 reveals the comprehensive medical tests and medical results of Student 1. The results in Tables 4 and 5 show the validation of the computational results as obtained in Fig. 3e and f.

4.1. Experimental evaluation of classification result

LIBSVM implements the SVM algorithm, which is an easy and efficient framework for solving medical image classification problems. The system implements the Radial Basis Function (RBF) kernel because of its simplicity and performance. We conducted an evaluation of the classification result. Six training image samples and 1 test image sample were collected from each subject, summing up to 7 sample images per subject. An accuracy of 91.7% was obtained from the classification which involved two subjects. A more detailed classification accuracy could be obtained with a larger training and test dataset (see Supporting documents 3 – FacialTrain.txt (saved as FacialTrain.doc), FacialTest.txt (saved as FacialTest.doc) and result.txt (result.doc)).

These depict the performance evaluation of the SVM classifier. The figure below (Fig. 3i) depicts the performance evaluation of the system. The scale of sensitivity was depicted on the y-axis, while that of specificity was depicted on the x-axis. Considering the abbreviations in the plot of Fig. 3i, the interpretations for the acronyms are, respectively, TPR (true-positive rate) versus FPR (false-positive rate). The Kappa statistics also produced an excellent result (0.9) [79] [see Tables 1–3].

ROC can be used to estimate algorithm performance [41]. Interpretation for ROC curves can be provided either graphically or numerically. The ROC Curve in Fig. 3i depicts a perfect probabilistic classifier. Such classifiers allocate higher scores to all the positives than to any of the negatives. In such a situation, the positives will appear on top of the ranked list. The ROC curve displayed has an infinitesimal change in the curve, which is easily missed because the TP, FP values displayed in that curve were respectively: TP = 1,0.5,1,1,1,1 FP = 0,0,0,0,0.1,0. The TP and FP values were obtained from the results generated by WEKA [a machine learning framework and software developed in Java] [80,81]. WEKA had a correctly classified instance of 11, which represents (91.6667%), and an incorrectly classified instance of 1 (8.3333%).

The evaluation metrics mentioned in Table 1 helped to depict the performance of the Support Vector Machine (SVM) used for the

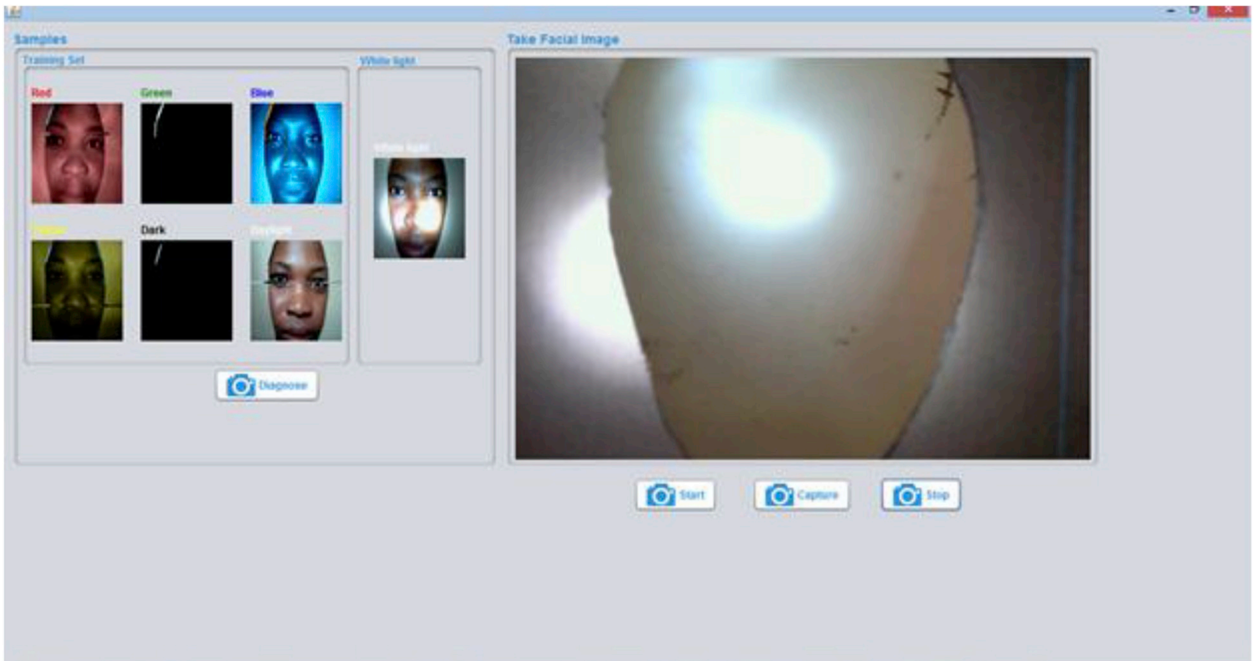


Fig. 3g. Facial image samples captured by the computational health informatics diagnostic software for student 2 in one of the Sub-Saharan universities.

classification purposes. The metrics are related to the proposed methodology because it used data collated during the Facial recognition experiments. The evaluation metrics are used for evaluating the performance of the SVM classifier in correctly classifying images collated from the experimental subjects. The results showed that 11 instances were correctly classified, while 1 was incorrectly classified. The next evaluation parameter gave a Kappa Statistic value of 0.9, an excellent result [79] (Viera and Garrett, 2005). The other evaluation parameters are the mean absolute error, root mean squared error, relative absolute error, and the root relative squared error. The values of these errors are respectively: 0.0278, 0.1667, 10%, and 44.7214%.

*Mean absolute error [82] is a type of error that is an unambiguous measure of average error magnitude. Mean absolute error has more advantages over root mean squared error in assessing model or algorithmic performances.

Table 2 provides information that helps to depict the detail accuracy of experimental evaluation of classification results, by revealing the values of the True Positive (TP) Rate, FR Rate, Precision, Recall, F-measure, ROC and area class. The color codes were also associated with the values depicted.

Table 3 presents the confusion matrix for evaluation of the classification result. The table enables the visualization of the algorithm performance used in classifying the results [83,84].

4.2. Validating the computational experimental results

In order to validate the computational results obtained from diagnosis conducted with our TCM-based computational health informatics

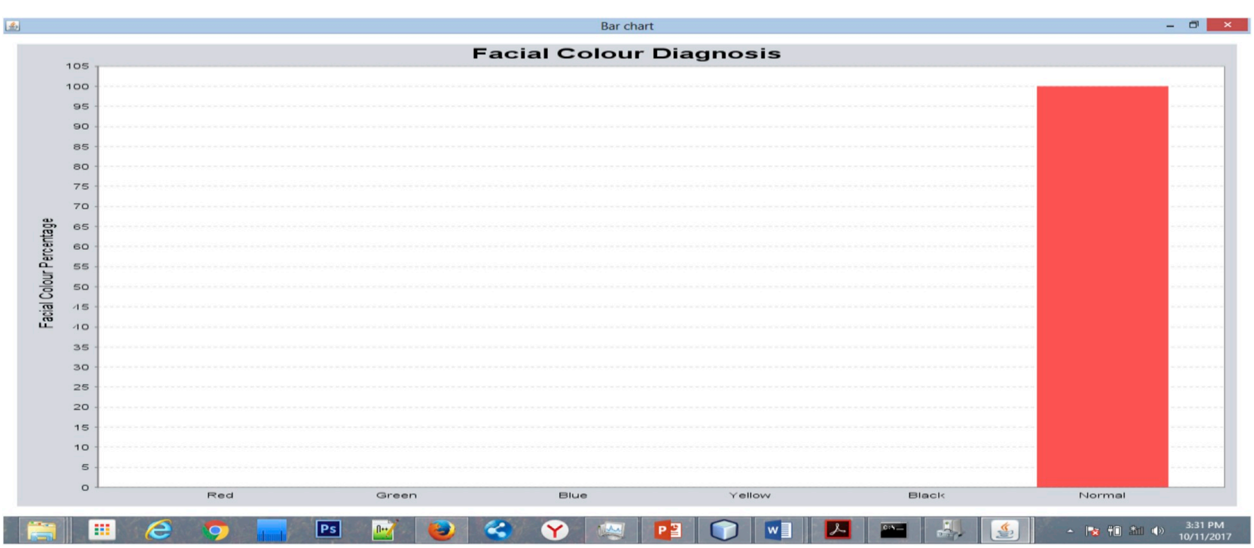


Fig. 3h. Facial Color Diagnosis Classification graph depicting output of the TCM-based computational health informatics diagnostic software for student 2. In TCM theory, five (5) internal organs; liver, lungs, spleen, heart, and kidney, are mapped to five colors; blue, red, yellow, white and black respectively, which may appear on the face. These colors show the conditions of these organs [13]. In our own results, blue, red, yellow, green, and black represent liver, lungs, spleen, heart and kidney respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

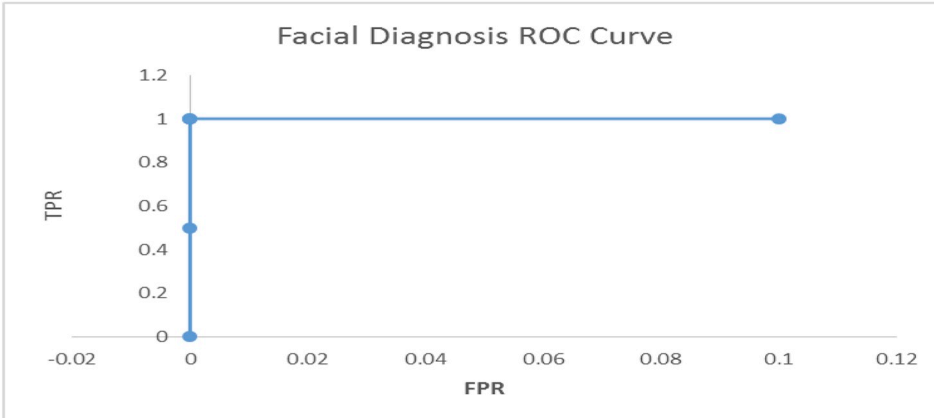


Fig. 3i. Facial diagnosis ROC Curve that plots TPR (True Positive Rate) versus FPR (False Positive Rate).

diagnostic tool, it was agreed that one of the experimental subjects (Student 1), whose face was captured by our TCM-based computational health diagnostic tool, should go for a hospital based medical evaluation in a standard health facility (Kogi State Specialist Hospital, Lokoja, Nigeria). The hospital has state-of-the-art medical facilities. Although the cost of conducting a comprehensive medical check-up in the hospital was expensive, as a result of this, only one of the patients (Student 1) could be funded to undergo the process. All the ethical documents relating to this were duly certified by the student and the medical hospital.

Table 4 shows the results obtained from our software versus the medical tests/check-up results.

5. Discussion

The result of the research produced the TCM-based computational health diagnostic software. The tool also has an associated hardware. The results highlighted in Fig. 3a, b, 3c, 3d, 3e, 3f, 3g, and 3h depict the output produced from our experimentation.

Fig. 3a and b shows the patient's registration form and the preliminary clinical examination data. On the data form, the clinical data icon will be clicked to collect patient clinical data.

Fig. 3c shows the contents of the record icon, which include the following: the search, update, insert, and delete diagnostic records. Fig. 3d depicts the diagnosis module, which consists of the "diagnose", "start", "capture", and "stop" buttons. It also has different colors as training sets, during the training and learning process of the SVM. It additionally has a section for displaying the facial image of the patient under examination.

Fig. 3e shows the diagnosis section of our software. This section depicts the facial image samples captured by the computational health diagnostic software for student 1. The student's facial image was acquired by the facial image acquisition chamber, captured, displayed, and trained under different colors of light.

Fig. 3f shows the output of the facial diagnosis. It depicts a facial color diagnosis classification graph. This clearly shows the output of the TCM-based computational health diagnostic software for student 1. Here, the red indicator points at the normal segment, which indicates that the patient is healthy. In TCM theory, five (5) internal organs; liver, lungs, spleen, heart, and kidney, are mapped to five colors; blue, red, yellow, white and black respectively, which may appear on the face. These colors show the conditions of these organs [13]. In the graphical results of our computational health diagnostic tool, blue, red, yellow, green, and black represent liver, lungs, spleen, heart and kidney respectively.

Fig. 3g shows the facial image samples captured by the computational health diagnostic software for student 2 in one of the Sub-Sahara universities within the diagnosis module of our software.

Fig. 3h shows the facial color diagnosis classification graph, depicting the output of the TCM-based computational health diagnostic software, for student 2. Here, the red indicator points at the normal segment which indicates that the patient (student 2) is healthy. In TCM theory, five (5) internal organs; liver, lungs, spleen, heart, and kidney, are mapped to five colors; blue, red, yellow, white and black respectively, which may appear on the face. These colors show the conditions of these organs [13]. In the graphical results of our computational health diagnostic tool, blue, red, yellow, green, and black depict liver, lungs, spleen, heart and kidney respectively.

Tables 4 and 5 show the validation of our computational experimental results with the medical check-up diagnostics conducted by medical personnel at a standard hospital. The medical results of student 1 was confirmed with the results predicted by our computational health informatics diagnostic software.

Fig. 3i shows the facial diagnosis ROC Curve that plots TPR (True Positive Rate) versus the FPR (False Positive Rate).

5.1. Further explanation of results with respect to the methodology guides on the evaluation of diagnostic tests

Based on the existing literature, we were able to examine the contents of our study and correlate the methods and results with the methodology guides on the evaluation of diagnostic tests. Some of the literature found [67,76,77], contained different metrics that can be used to evaluate diagnostic tests. Some of the literature recommended performance of medical tests as one of the ways to validate diagnosis. Some other parameters for validating diagnosis include sensitivity and specificity measurements. One of the literature studies [77] highlighted that diagnostic accuracy studies would require certain parameters for validation. Such parameters include: medical tests, intended use of the test, specificity, sensitivity, role of the test, and target condition, amongst others. We conducted a medical test to validate the experimental results of one of the experimental subjects (Student 1). We

Table 1
Statistical summary of experimental evaluation of classification result.

Evaluation Summary		
Correctly Classified Instances	11	91.6667%
Incorrectly Classified Instances	1	8.3333%
Kappa statistic	0.9	
Mean absolute error	0.0278	
Root mean squared error	0.1667	
Relative absolute error	10%	
Root relative squared error	44.7214%	
Total Number of Instances	12	

Table 2
Detail accuracy of experimental evaluation of classification result.

Detailed Accuracy By Class							
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC	Area Class
	1	0	1	1	1	1	Red
	0.5	0	1	0.5	0.667	0.75	Green
	1	0	1	1	1	1	Blue
	1	0	1	1	1	1	Yellow
	1	0.1	0.667	1	0.8	0.95	Black
	1	0	1	1	1	1	Normal
Weighted Avg.	0.917	0.017	0.944	0.917	0.911	0.95	

Table 3
Confusion matrix of Experimental Evaluation of Classification Result.

Confusion Matrix						
a	b	c	d	e	f	< – classified as
2	0	0	0	0	0	a = red
0	1	0	0	1	0	b = green
0	0	2	0	0	0	c = blue
0	0	0	2	0	0	d = yellow
0	0	0	0	2	0	e = black
0	0	0	0	0	2	f = normal

Table 4
Results predicted by the Computational Health Diagnostic Software Versus Medical Test Result for. Student 1.

Student 1	Computational Health Diagnostic Software Result	Comprehensive Medical Test Result
Liver	Normal (100%)	100%Normal
Lungs	Normal (100%)	99% Normal
Spleen	Normal (100%)	100%Normal
Heart	Normal (100%)	100%Normal
Kidney	Normal (100%)	100%Normal
Result Summary	100% Normal	Normal and Certified Okay.

further computed the specificity, sensitivity of the classification obtained from the computational experimental result. In our results, the experimental subjects were healthy. Validated medical tests on one of the experimental subjects also showed that the subject was healthy. This result shows that the TCM-based computational health diagnostic tool can be used for disease diagnosis, and to identify healthy students.

5.2. Significance of research

The TCM-based facial image diagnostic computational health diagnostic software can provide an efficient, effective, and reliable way to easily conduct screening and analysis of the health status of students, in different university health centers, within the Sub-Saharan African region. It will particularly complement the efforts of medical doctors and medical personnel in the area of disease diagnosis, by providing prompt diagnosis, which saves time that would have been spent on undergoing a full medical check-up. The prospect of such prompt diagnosis can assist in early detection of non-communicable diseases.

This will encourage people to engage in the computational health diagnostic software test. Furthermore, our software will help provide an inexpensive and affordable quality healthcare diagnosis facility to students in higher institutions, within the Sub-Saharan African region, thus making regular medical check-up among students consistent. It will also help enhance a good state of health among the youths, and prevent untimely deaths among undergraduate and postgraduate students in higher institutions, within the Sub-Saharan African region. Medical doctors cannot diagnose a patient by merely looking at the patient's face. The software will be vital during emergencies when mortality rates can be greatly reduced by means of early detection of a disease. The identity of patients can be easily verified by the computational health diagnostic software, and faster diagnosis can be achieved, which helps medical personnel to commence the first set of treatments, thus saving the lives of the students. This can also help avert a crisis on campus. More so, the software will increase the productivity of medical experts. It will help to store medical data and record in a secure database as a secondary backup mechanism. It has a helpful User Interface (UI), which will enhance user friendliness and interaction. Finally, this is probably the first-of-its kind, as applied on an African population. Further work is still needed. More data is needed to further test the computational health informatics diagnostic tool.

5.3. Limitations

Some of the limitations that we encountered in the course of this research include: lack of access to medical data from hospitals. Different hospitals were reluctant and unwilling to release patients' medical report and data even without identifiers. Many students who engaged in smoking and alcohol consumption, as a lifestyle, in different Nigerian universities, were unwilling to present themselves as experimental subjects for our computational health informatics diagnostic tool and for medical check-up/screening. Moreover, some of the TCM papers were written in the Chinese language. This is one of the limitations associated with the systematic literature review of this study. In order to validate results predicted by our TCM-based computational health informatics diagnostic software, the cost of conducting medical checks/tests for experimental patients was expensive. Conducting medical tests also took some time for the results to be ready.

6. Conclusion

The implementation of a TCM-based computational health informatics diagnostic software, through facial diagnosis, is a step in the right direction, and will greatly benefit students in higher institutions, within the Sub-Saharan African region. It will be good if the Government of nations within Sub-Saharan Africa, can properly invest in this technology, and provide health centers of different higher institutions with this computational tool, for effective, efficient diagnosis and treatment of students.

Supplementary materials: (Supplementary material 1 and Supplementary material 2), contain the supporting documents for this research paper– Files, letters, ethical consent, medical test results, excel; files, figures, tables, amongst others.

Ethics and consent to participate

Written consent for participating in the study and undergoing TCM-based facial diagnosis, and medical examinations, was obtained from each participant prior to the experiments.

The consent of the Kogi State Specialist Hospital was obtained about publishing the results of the medical tests for research and publication purposes.

Table 5
Results obtained from comprehensive medical test physically conducted for Student 1 at the State Specialist Hospital, Lokoja, Nigeria.

	Test1/Result	Test 2/Result	Test 3/Result	Test 4/Result	Test 5/Result	Outcome
Liver	Alanine Transaminase (ALT) Test:1.2 U/L	Aspartate Transaminase (AST) Test:1.9 IU/L	Alkaline Phosphatase (ALP) Test:	Albumin Test:25 g/L (Acceptable Range: 35–55 g/L)	Bilirubin Test:13uMol/L Normal acceptable range: 2-20uMol/L; Source: [78],https://www.healthinfo.org.nz/patientinfo/269153.pdf	Normal
Lungs	a blood oxygen level test done. A probe is placed on the tip of your finger and the amount of oxygen you are breathing in is measured.	`SpO2 @1 mm is 99%; PR-72bpm				Normal
Spleen	Blood tests, such as a complete blood count to check the number of red blood cells, white blood cells and platelets in your system	WBC: 5.1 Plat: 33.7 Pvc:42% Hb/g/dl:14 NEUT:57.4% LYMPH MONO:35.9% EOSIN:6.7%				Normal
Heart	Checking Your Blood Pressure: 120/80 (100%)	a blood test to check your levels of sodium, potassium, albumin, and creatinine. Abnormal levels could suggest problems with organs like your kidneys and liver, possible signs of heart failure.	Sodium:138 mmol/L (Acceptable Range: 135–145 mmol/L) Potassium: 3.8 mmol/L (Acceptable Range: 3.5–5.0 mmol/L) Albumin: Test:25 g/L (Acceptable Range: 35–55 g/L) Creatinine: 84 µmol/L (Acceptable Range: 60-130 µmol/L)			Normal
Kidney	ACR (Albumin to Creatinine Ratio) your urine.	ACR stands for “albumin-to-creatinine ratio	Albumin: Test:25 g/L/Creatinine: 84 µmol/L = 0.2976 (Normal)			Normal

Author contributorship

OO conceived and designed the experiments. OO and AJ performed the experiments. OO and AJ analyzed the results. OO wrote the first draft of the manuscript. OO wrote the revised version of the manuscript. All authors read and approved the final manuscript.

Conflict of interest declaration

The authors declare that there are no competing interests.

Acknowledgement

We appreciate The Oluwa Research, Development and Philanthropic Foundation (TORDPF), Nigeria, for providing resources (computers, internet, hospital medical tests and screening funds, etc.), at the commencement of this research.

The later part of the research was partly supported by the DST/NRF Innovation Postdoctoral Fellowship Award funding, South Africa, with (Grant no: UID: 111988), awarded to Dr. Olugbenga Oluwagbemi.

The grant holder acknowledges that opinions, findings, and conclusions or recommendations expressed in this publication of this NRF supported research, are those of the authors and that the NRF accepts no liability whatsoever in this regard.

We thank Olaoluwa Bamiduro for designing the graphical abstract.

We thank Olaoluwa DuroBello, of the University of Nottingham, Malaysia, for taking time to correct the grammar and English language constructs of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.imu.2018.12.002.

References

[1] Teo CH, Ng CJ, White A. What do men want from a health screening mobile app? A qualitative study. PLoS One 2017;12(1). e0169435 https://doi.org/10.1371/journal.pone.0169435.
[2] Culica D, Rohrer J, Ward M, Hilsenrath P, Pomrehn P. Medical checkups: who does not get them? Am J Public Health 2002;92(1):88–91. PMID:11772768.
[3] Dryden R, Williams B, McCowan C, Themessl-Huber M. What do we know about who does and does not attend general health checks? Findings from a narrative scoping review. BMC Public Health 2012;12:723. PMID: 22938046.
[4] Hoebel J, Starker A, Jordan S, Richter M, Lampert T. Determinants of health check attendance in adults: findings from the cross-sectional German Health Update (GEDA) study. BMC Public Health 2014;14. 913 PMID: 25185681.
[5] Mao L, de Wit JB, Kippax SC, Prestage G, Holt M. Younger age, recent HIV diagnosis, no welfare support and no annual sexually transmissible infection screening are associated with nonuse of antiretroviral therapy among HIV-positive gay men in Australia. HIV Med 2015;16(1):32–7. PMID: 24889053.
[6] Reeder AI. It's a small price to pay for life: faecal occult blood test (FOBT) screening for colorectal cancer, perceived barriers and facilitators. N Z Med J 2011;124(1331):11–7. PMID: 21725408.
[7] Vincent J, Hochhalter AK, Broglio K, Avots-Avotins AE. Survey respondents planning to have screening colonoscopy report unique barriers. Perm J 2011;15(1):4–11. PMID:2150561.
[8] Palmer RC, Midgette LA, Dankwa I. Colorectal cancer screening and African Americans: findings from a qualitative study. Cancer Control 2008;15(1):72–9. PMID:18094663.
[9] http://www.who.int/gho/ncd/en/; Access date December 18th, 2017.
[10] Li X, Li F, Wang Y, Quian P, Zheng X. Computer-aided disease diagnosis system in TCM based on facial image analysis. Int J Funct Inf Personalised Med (IJFIPM) 2009;2(3):303–14.
[11] Li F, Zhou R, Li G, Zhao R. Classification of facial diagnosis gloss in Chinese medicine based on different algorithms. Chin J Integr Med 2016:1–6.
[12] Li FuFeng, Zhao Changbo, Zheng Xia, Wang Yiqin, Zhou Xiaobo, Li Guo-Zheng. Computer-assisted lip diagnosis on traditional Chinese medicine using multi-class support vector machines. BMC Complement Altern Med 2012;12:127. https://doi.org/10.1186/1472-6882-12-127.
[13] Fufeng L, Dan D, Xiaoqiang L, Yiqin W, Peng Q, Xiaoyan Z, Guoping L. Facial complexion acquisition and recognition system for clinical diagnosis in Traditional Chinese Medicine. International joint conference on bioinformatics, systems biology and intelligent computing. IEEE Computer Society; 2009. p. 1–5.
[14] Zhao C, Li G, Li F, Wang Z, Liu C. Qualitative and quantitative analysis for facial complexion in traditional Chinese medicine. BioMed Res Int 2014;2014. Article ID 207589, 17 pages https://doi.org/10.1155/2014/207589.
[15] Zhang J, Hou S, Wang J, Li L, Li P, Han J, Yao H, Sun R, Li Z, Lei Z, Wang Q. Classification of traditional Chinese medicine constitution based on facial features in color images. Journal of Traditional Chinese Medical Sciences 2016;3(3):141–6.
[16] Luo C-H, Ye J-W, Lin C-Y, Lee T-L, Tsai L-M, Shieh M-D. L-cube polynomial for the recognition of normal and hypertensive string-like pulse mappings in Chinese medicine. Informatics in Medicine Unlocked In press 2018;12:27–33https://doi.org/10.1016/j.imu.2018.05.006.
[17] Yang Y, Zhang J, Zhuo L, Cai Y , Zhang X.(2012). Cheek region extraction method for face diagnosis of Traditional Chinese Medicine, 2012 IEEE 11th international conference on signal processing (ICSP), held in beijing, China between 21st -25th october, 2012.
[18] Liu G-P, Li G-Z, Wang Y-L, Wang Y-Q. Modelling of inquiry diagnosis for coronary heart disease in traditional Chinese medicine by using multi-label learning. BMC Complement Altern Med 2010;10:37. 2010 http://www.biomedcentral.com/1472-6882/10/37.
[19] Mist S, Ritenbaugh C, Aickin M. Effects of questionnaire-based diagnosis and training on inter-rater reliability among practitioners of traditional Chinese medicine. J Alternative Compl Med 2009;15(7):703–9.
[20] Lo LC, Cheng TL, Chen YJ, Natsagdorj S, Chiang JY. TCM tongue diagnosis index of early-stage breast cancer. Complement Ther Med 2015;23(5):705–13. https://doi.org/10.1016/j.ctim.2015.07.001. Epub 2015 Jul 10.
[21] Kang H, Zhao Y, Li C, Chen Y, Tang K, Yang L, Ma C, Peng J, Zhu R, Qi Liu, Hu Y, Cao Z. Integrating clinical indexes into four-diagnostic information contributes to the traditional Chinese medicine (TCM) syndrome diagnosis of chronic hepatitis B. 2015 Sci Rep 2015;5:9395. https://doi.org/10.1038/srep09395. Published online 2015 Mar 23.
[22] Xue D, Han S, Jiang S, Sun H, Chen Y, Li Y, Wang W, Feng Y, Wang K, Li P. Comprehensive geriatric assessment and traditional Chinese medicine intervention benefit symptom control in elderly patients with advanced non-small cell lung cancer. Med Oncol 2015;32:114. 2015.
[23] Chen Y, Liu W, Zhang L, Yan M, Zeng Y. Hybrid facial image feature extraction and recognition for non-invasive chronic fatigue syndrome diagnosis. Comput Biol Med 2015;64:30–9. https://doi.org/10.1016/j.compbiomed.2015.06.005.
[24] Li G-Z, He Z, Shao F-F, Ou A-H, Lin X-Z. Patient classification of hypertension in Traditional Chinese Medicine using multi-label learning techniques 2015 BMC Med Genomics 2015;8(Suppl 3). S4.
[25] Watsuji T, Arita S, Shinohara S, Kitade T. Medical Application of fuzzy theory to the diagnostic system of tongue inspection in traditional Chinese medicine, 1999. IEEE international fuzzy systems conference proceedings august 22-25, 1999, seoul, korea. 1999.
[26] Moura NG, Cordovil I, Ferreira Ade S. Traditional Chinese medicine wrist pulse-taking is associated with pulse waveform analysis and hemodynamics in hypertension. Journal of Integrative Medicine 2016;14(2):100–13. https://doi.org/10.1016/S2095-4964(16)60233-9.
[27] Jiang RS, Liang KL. Establishment of olfactory diagnosis for the traditional Chinese version of the university of Pennsylvania smell identification test. 2016 Dec Int Forum Allergy Rhinol 2016;6(12):1308–14. https://doi.org/10.1002/alr.21818. Epub 2016 Jun 22.
[28] Xu J, Xu ZX, Lu P, Guo R, Yan HX, Xu WJ, Wang YQ, Xia CM. Classifying syndromes in Chinese medicine using multi-label learning algorithm with relevant features for each label. Chin J Integr Med 2016;22(11):867–71. Epub 2016 Oct 26.
[29] Wu HK, Ko YS, Lin YS, Wu HT, Tsai TH, Chang HH. The correlation between pulse diagnosis and constitution identification in traditional Chinese medicine. 2017 Feb Complement Ther Med 2017;30:107–12. https://doi.org/10.1016/j.ctim.2016.12.005.
[30] Zheng S, Kim C, Meier P, Sibbritt D, Zaslawski C. Development of a novel questionnaire for the traditional Chinese medicine pattern diagnosis of stress. Journal of Acupuncture and Meridian Studies 2017;10(4):276–85. https://doi.org/10.1016/j.jams.2017.06.002.
[31] Zhang J, Xu J, Hu X, Chen Q, Tu L, Huang J, Cui J. Diagnostic method of diabetes based on support vector machine and tongue images. BioMed Res Int 2017;2017:7961494. 9 pages https://doi.org/10.1155/2017/7961494.
[32] Tian X, Zhu G, Wang J, Wang Q, Guan L, Tan Y, Xue Z, Qin L, Zhang J. Study on the relation between tissues pathologies and traditional Chinese medicine syndromes in knee osteoarthritis: medical image diagnostics by preoperative X-ray and surgical arthroscopy. J X Ray Sci Technol 2016;24(4):509–19. https://doi.org/10.3233/XST-160567. 2016 Apr 7.
[33] Jiang B, Liang X, Chen Y, Ma T, Liu L, Li J, Jiang R, Chen T, Zhang X, Li S. Integrating next-generation sequencing and traditional tongue diagnosis to determine tongue coating microbiome. Sci Rep 2012;2:936. https://doi.org/10.1038/srep00936. 2012 Epub 2012 Dec 6.
[34] Wang H, Cheng Y. A quantitative system for pulse diagnosis in Traditional Chinese Medicine. IEEE Conference Proceeding Eng Med Biol Soc 2005;6:5676–9. 2005.
[35] Chiu CC, Chang HH, Yang CH. Objective auscultation for traditional Chinese medical diagnosis using novel acoustic parameters. Comput Methods Progr Biomed 2000;62(2):99–107.
[36] O'Brien KA, Abbas E, Zhang J, Guo ZX, Luo R, Bensoussan A, Komesaroff PA. Understanding the reliability of diagnostic variables in a Chinese Medicine examination. J Alternative Compl Med 2009;15(7):727–34.
[37] Hua B, Abbas E, Hayes A, Ryan P, Nelson L, O'Brien K. Reliability of Chinese medicine diagnostic variables in the examination of patients with osteoarthritis of the knee. J Alternative Compl Med 2012;18(11):1028–37.
[38] Ferreira A. Diagnostic accuracy of pattern differentiation algorithm based on Chinese medicine theory: a stochastic simulation study 2009 Chin Med 2009;4. 24.

[39] Ferreira AS. Misdiagnosis and undiagnosis due to pattern similarity in Chinese medicine: a stochastic simulation study using pattern differentiation algorithm. Chin Med 2011;6:1. 2011.
[40] Jeon YJ, Kim JU, Lee HJ, Lee J, Ryu HH, Lee YJ, Kim JY. A clinical study of the pulse wave characteristics at the three pulse diagnosis positions of Chon. Gwan and Cheok, Evidence-Based Complementary and Alternative Medicine 2011:9. 2011, Article ID 904056.
[41] Hui SC, He Y, Thach DTC. Machine learning for tongue diagnosis, 2007. 6th IEEE international conference on information, communications & signal processing, held in Singapore, between 10-13 dec. 2007. 2007.
[42] Yuen PC, Kuang ZY, Wu W, Wu YT. Tongue texture analysis using Gabor Wavelet opponent color features for tongue diagnosis in traditional Chinese medicine, (2000). Texture Analysis in Machine Vision 2000;40:179–88.
[43] Bakshi D, Pal S. Introduction about traditional Tongue Diagnosis with scientific value addition. 2010 IEEE international conference on systems in medicine and biology, held in India, between 16-18 dec. 2010 2010. https://doi.org/10.1109/ICSMB.2010.5735385.
[44] Hu C-S, Chung Y-F, Luo C-H, Yeh C-C, Si X. Pulse differences and 3D pulse mapping in TPNI displacements. 2011 4th international conference on biomedical engineering and informatics (BMEI), held in China, between, 15th to 17th, , october, 2011. 2011.
[45] Anastasi JK, Chang M, MS, Quinn J, Capili B. Tongue inspection in TCM: observations in a study sample of patients living with HIV. Med Acupunct 2014;26(1):15–22. https://doi.org/10.1089/acu.2013.1011. 2014 Feb 1.
[46] Zhao Y, Kang H, Peng JH, Xu L, Cao ZW, Hu YY. Key symptoms selection for two major syndromes diagnosis of Chinese medicine in chronic hepatitis B. 2017 Apr Chinese Journal of Integrated Medicine 2017;23(4):253–60. https://doi.org/10.1007/s11655-016-2253-3. Epub 2016 May 25.
[47] Oluwagbemi O, Oluwagbemi F, Abimbola O. *Ebinformatics:* Ebola fuzzy informatics systems on the diagnosis, prediction and recommendation of appropriate treatments for Ebola virus disease (EVD). Informatics in Medicine Unlocked2 2016:12–37. 2016.
[48] Oluwagbemi O, Oluwagbemi F, Ughamadu C. Android mobile informatics application for some hereditary diseases and disorders (AMAHD): a complementary framework for medical practitioners and patients. Informatics in Medicine Unlocked 2016;2:38–69. 2016.
[49] Oluwagbemi O, Oladunni B. Diagnosis and recommender system for some neglected tropical diseases. Int J Nat Appl Sci 2010;6(2):181–8. 2010.
[50] Oluwagbemi OO, Ofoezie U, Nwinyi O. A Knowledge-based data mining system for diagnosing malaria related cases in Healthcare Management. Egyptian Computer Science Journal 2010;34(4):23–34.
[51] Oluwagbemi O, Adeoye O, Fatumo S. Building a computer-based expert system for malaria environmental diagnosis: an alternative malaria control strategy. Egyptian Computer Science Journal 2009;33(1):55–69.
[52] Oluwagbemi Olugbenga O, Oluwagbemi Folakemi E, Fagbore Olatunji. *Malavefes*: a computational voice-enabled malaria fuzzy informatics software for correct dosage prescription of anti-malaria drugs. Journal of King Saud University – Computer and Information Sciences 2018;30(2):185–97.
[53] Heron S. Advanced encryption standard (AES). Netw Secur 2009;2009(12):8–12.
[54] Dworkin MJ, Barker EB, Nechvatal JR, James Foti J, Bassham LE, Roback E, Dray JF. Advanced encryption standard (AES). Federal inf. Process. Stds. (NIST FIPS) vol. 197. 2001.
[55] Nechvatal J, Barker E, Bassham L, Burr W, Dworkin M, Foti J, Roback E. Report on the development of the advanced encryption standard (AES). Journal of Research of the National Institute of Standards and Technology 2001;106(3):511–77.
[56] Lu C-C, Tseng S-Y. Integrated design of AES (advanced encryption standard) encrypter and decrypter, the IEEE international conference on application-specific systems, architectures and processors. 2002. Proceedings, held at san jose, CA, USA, between 17th to 19th july, 2002. 2002.
[57] Morkel T, Eloff JHP, Olivier MS. An overview of image steganography. The proceedings of the fifth annual information security South Africa conference. 2005.
[58] Li B, He Junhui, Huang Jiwu, Shi Yun Qing. A survey on image steganography and steganalysis. Journal of Information Hiding and Multimedia Signal Processing 2011;2(2):142–72.
[59] Kharrazi M, Sencar HT, Memon N. Image steganography and steganalysis: concepts and practice. Mathematics and computation in imaging science and information processing. https://doi.org/10.1142/9789812709066_0005; 2007. 177-207.
[60] Chaves-González JM, Vega-Rodríguez MA, Gómez-Pulido JA, Sánchez-Pérez JM. Detecting skin in face recognition systems: a color spaces study. Digit Signal Process 2010;20(3):806–23.
[61] Wang X, Zhang B, Guo Z, Zhang D. Facial image medical analysis system using quantitative chromatic feature. Expert Syst Appl 2013;40(9):3738–46.
[62] Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics 1973;3(6):610–21.
[63] ColorSummarizer: (Martin Krzywinski). http://mkweb.bcgsc.ca/color-summarizer/?api -, Accessed date: 18 December 2017.
[64] Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 2012;9(7):671–5. PMID 22930834.
[65] Schindelin J, Rueden CT, Hiner MC, Eliceiri KW. The ImageJ ecosystem: an open platform for biomedical image analysis. Mol Reprod Dev 2015;82(7-8):518–29. https://doi.org/10.1002/mrd.22489. Epub 2015 Jul 7, PMID 26153368.
[66] Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, Eliceiri KW. ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinf 2017;18. https://doi.org/10.1186/s12859-017-1934-z. 529.
[67] Gopalakrishna G, Mustafa RA, Davenport C, Scholten RJ, Hyde C, Brozek J, Schünemann HJ, Bossuyt PM, Leeflang MM, Langendam MW. Applying grading of recommendations assessment, development and evaluation (GRADE) to diagnostic tests was challenging but doable. J Clin Epidemiol 2014;67(7):760–8. https://doi.org/10.1016/j.jclinepi.2014.01.006.
[68] Image: https://imagej.net/Welcome- Accessed 18 December 2017.
[69] Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) 2011;2(3). https://doi.org/10.1145/1961189.1961199.
[70] Chang, Lin. LIBSVM: version 3.22. https://www.csie.ntu.edu.tw/~cjlin/libsvm/; 2016, Accessed date: 18 December 2017.
[71] Wang Y-Q. An analysis of the viola-jones face detection algorithm. Image Process Line 2014;4:128–48. 2014 https://doi.org/10.5201/ipol.2014.104.
[72] Connolly C, Fleiss T. A study of efficiency and accuracy in the transformation from RGB to CIELAB color space. IEEE Trans Image Process 1997;6(7):1046–8.
[73] Nishad PM, Chezian RM. Various color spaces and color space conversion algorithms. Journal of Global Research in Computer Science 2013;4(1):44–8.
[74] Haralick RM, Shamugam K, Dinstein I-H. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics (SMC) 1973;3(6):610–21.
[75] Liu K, Cheng Y-Q, Yang J-Y. Algebraic feature extraction for image recognition based on an optimal discriminant criterion. Pattern Recogn 1993;26(6):903–11.
[76] Methods guides for medical review tests", Source: agency for healthcare research and quality. Methods guide for medical test reviews [posted november 2010]. Rockville, MD. http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-andreports/?pageaction=displayProduct&productID=454, Accessed date: 10 February 2018.
[77] Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine D, Reitsma JB, de Vet HCW, Bossuyt PMM. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. 2016 BMJ Open 2016;6. https://doi.org/10.1136/bmjopen-2016-012799. e012799.
[78] https://www.healthinfo.org.nz/patientinfo/269153.pdf.
[79] Viera AJ, Garrett JM. Understanding interobserver agreement: the Kappa statistic. Fam Med 2005;37(5):360–3.
[80] Holmes G, Donkin A, Witten IH. Weka: a machine learning workbench. Proceedings of the Second Australia and New Zealand conference on intelligent information systems, brisbane, Australia. 1994.
[81] Witten IH, Frank E, Trigg L, Hall M, Holmes G, Jo Cunningham S. Weka: practical machine learning tools and techniques with java implementations. Proceedings of the ICONIP/ANZIIS/ANNES'99 workshop on emerging knowledge engineering and connectionist-based information systems. 1999. p. 192–6. Retrieved 2007-06-26.
[82] Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim Res 2005;30:79–82.
[83] Stehman SV. Selecting and interpreting measures of thematic classification accuracy. Rem Sens Environ 1997;62(1):77–89. https://doi.org/10.1016/S0034-4257(97)00083-7.
[84] Powers DM. Evaluation: from precision, Recall and F-measure to ROC, informedness, markedness & correlation. J Mach Learn Technol 2011;2(1):37–63.