The Adoption of Modified Convolutional Neural Networks Algorithm Based Authentication Scheme for Examination Conduct in Nigeria

Siyaka, H.O., Owolabi, O. & Hashim, B.I.
Department of Computer Science
University of Abuja
Abuja, FCT, Nigeria
E-mail: alhassan77077@gmail.com
Phone: +2348036331160

ABSTRACT

There is the need to urgently improve on the obtainable levels of quality of educational system and practice in Nigeria. The area of examination conducts leaves much to be desired due to prevalent malpractices and irregularities such as impersonation of candidates, extraneous materials smuggling into venues of examinations and poor authentication systems. The biometric technology came about to solve the problems of impersonation of candidates in which fingerprints were used to authenticate users during UTME in recent years. However, the quality of services was highly ineffective in terms of timing, speed and accuracy. At present, the face recognition evolved to make use of main geometrical features of the face of persons such as eye, mouth, and nose. This paper highlights a new model for autonomous face recognition based on modified CNN with pictorial and geometry characteristics of candidate's faces sitting for UTME in Nigeria. Therefore, the practical implementation of proposed model is to be carried out in the future works in order to ascertain the accuracy, speed and processing time.

Keywords: Biometrics, Face Authentication, Face, Recognition, CNN, Accuracy, Speed, SVM, UTME Processing Time, Features Extraction.

23rd iSTEAMS Conference Proceedings Reference Format

Siyaka, H.O., Owolabi, O. & Hashim, B.I. (2020): The Adoption of Modified Convolutional Neural Networks Algorithm Based Authentication Scheme for Examination Conduct in Nigeria. Proceedings of the 23rd iSTEAMS Conference, American University of Nigeria, Yola. April, 2020. Pp 143-156
www.isteams.net/yola2020

1. INTRODUCTION

In 1960's the first semi-automated facial recognition programs were created by Woody Bledsoe, Helen Chan Wolf, and Charles Bisson. These programs required the administrator to locate features such as the eyes, ears, nose, and mouth on the photograph, then calculate distances and ratios to a common reference point or reference data. Face recognition is a part of a wide area of pattern recognition technology. Face identification is the process of identifying a person in a digital image or video, and showing their authentication identity. Identification is a one-to-many matching process that compares a query face image against all the template images inside the face database in order to determine the identity of the query face. Identification mode allows both positive and negative recognition outcomes, but the results are much more computationally costly if the template database is large (Jones, 2009; Mau, Dadgostar & Lovell, 2012). In general, face recognition requires a set of visual tasks to be performed robustly. That process includes mainly three-task acquisition, normalisation and recognition.

The term acquisition means the detection and tracking of face-like image patches in a dynamic scene. Normalisation is the segmentation, alignment and normalisation of the face images, and finally recognition that is the representation and modelling of face images as identities, and the association of novel face images with known models.

A face is a typical multidimensional structure and needs good computational analysis for recognition. The overall problem is to be able to accurately recognize a person's identity and take some actions based on the outcome of the recognition process. Recognizing a person's identity is important mainly for security reasons, but it could also be used to obtain quick access to medical, criminal, or any type of records. Solving this problem is important because it could allow people to take preventive action, provide better service in the case of a doctor appointment, and allow users access to a secure area.

Face recognition is one of the few biometric methods that possess the merits of both high accuracy and low intrusiveness. It has the accuracy of a physiological approach without being obtrusive. Over past few decades, many face recognition techniques have been proposed, motivated by the increased number of real world applications requiring human face recognition. The importance of automatic face recognition is much better with numerous variations of images of the same face due to changes in the parameters such as pose, illumination, expression, motion, facial hair, glasses, and background (Sagar, & Narasimha, 2019).

Examination malpractice has been bedevilling the quality of graduates from different institutions in Nigeria and other parts of the world. It ranges from bringing extraneous material into examination halls to employing machineries that stand in place of candidates to write exam for them. Now with the advent of conducting and processing of examination result with the computer, there is a dire need of a system that is capable of authenticating the candidate that seat down for the examination. This paper investigates the process of utilizing Convolutional Neural Networks (CNNs) for the purpose of authenticating candidates seating for Unified Tertiary Matriculation Examination (UTME).

2. LITERATURE REVIEW

2.1 Face Recognition with CNN

Convolutional Neural Networks (CNN) is one of the most innovative solutions of the past 10 years in the application areas of pattern recognition, object detection and image recognition. It evolved from traditional Artificial Neural Networks in which the parameters quantity in the later is minimised. This makes CNN to perform complex computational problems hitherto unresolved by Classic ANN. The face recognition application of CNN takes into account images without recourse to position or location of faces. It is capable of extracting abstract features whenever input moves towards the deeper layers. During image classification problem, the edge could be detected in the first layer; the second layer could determine the simpler shapes, and the higher level of features such as faces in the proceeding layers (Albawi & Mohammed, 2017). The method searches for local regions in images as against whole image for effective classification and recognition tasks.

2.1.2 Advantages of Applying CNN for Face Recognition

Convolutional Neural Network has had ground breaking results over the past decade in a variety of fields related to pattern recognition; from image processing to voice recognition. The most beneficial aspect of CNNs is reducing the number of parameters in ANN. This achievement has prompted both researchers and developers to approach larger models in order to solve complex tasks, which was not possible with classic ANNs (Albawi & Mohammed, 2017). The most important assumption about problems that are solved by CNN should not have features which are spatially dependent. In other words, in a face detection application, attention is not paid to where the faces are located in the images. The only concern is to detect them regardless of their position in the given images (Bernal et al., 2018).

Another important aspect of CNN is to obtain abstract features when input propagates toward the deeper layers. In image classification, the edge might be detected in the first layers, and then the simpler shapes in the second layers, and then the higher level features such as faces in the next layers. A standard convolution neural network for image classification is made up of sequence of convolution layers, pooling layers and full-connection layers (Zhao, Lang, & Li, 2019).

In 1998, a network architecture known as Convolutional neural network is proposed by Yann LeCun and others. Previously, a number of works have been carried out in deep leaning and computer vision. Consequently, a number of research ventures evolved including image recognition, target detection, image classification (Becherer, Pecarina, & Hopkinson, 2019), speech recognition, character and phoneme recognition (Hwang & Sung, 2013).

The most important contribution of CNN network is the convolution in which non-linearity and pooling layer appended such as AlexNet and LeNet. In CNNs, weight sharing gives invariance translations to the model. It assists the filter to learn features regardless of the spatial properties. It uses the concept of pooling (or down-sampling) in order to reduce the complexity further layers. It is most appropriate for image processing domain, which minimising resolution. In the same vain, dropout approach is used to eliminate nodes and connections for the purpose of reducing complex computation and parameters. This improves the performance of the network due to smaller levels of layer requiring faster training and testing time. CNN is term most powerful machine learning technique for several field of applications including face and image detection, video and voice recognitions (Albawi & Mohammed, 2017).

2.1.3 Features of Convolutional Neural Networks

In present-day industries, surveillance networks have become significant component for many applications such as monitoring of patients in hospitals, detecting stadiums violence, and identifying misplaced luggage in airports. Notwithstanding the domain of application, one underlying objective is to observe and report interactions of interest. In addition, human interface is needed to monitor numerous camera feeds and react promptly as the occasion deserves. However, the increasing quantity of cameras in surveillance systems is present-day concerns for a single human to monitor because of enormous data streams almost immediately. Again, cost implication for number of camera feeds is undesirable and susceptible to errors (Cameron, Savoie, Kaye, & Scheme, 2019).

Consequently, the research focus shifted towards the advancement of computer-based intelligent systems for monitoring surveillance camera systems. By applying the principle of machine learning principles and contemporary computer vision algorithms; it became feasible to accurately detect objects of interest in an image frame, and classify them surpassing accuracy offered by humans (Sze, Chen, Yang, & Emer, 2017). Convolutional neural networks (CNNs) are a top choice in detecting and categorizing digital images (LeCun, Bengio, & Hinton, 2015). In fact, these networks take advantage of its deep structure made up of convolutional layers, pooling layers, fully connected layers, and non-linear activation functions.

The typical CNNs structure converts the input data into a deep low-dimensional feature format in order to enhance the classification task undertaken by the fully connected layers (Cameron et al., 2019). This work is to investigate CNN in the authenticating students for Unified Tertiary Matriculation Examination using their acquired images during registration stages for speed and better accuracy.

2.2 The Concept of Biometric System

Soutar, Roberge, Stoianov, Rene, & Kumar (1999) defined biometrics as a unique, measurable, biological characteristic or trait for automatic recognition or verification of the identity of a human being. A person's individuality can be differentiated from one or more behavioural or physiological features by biometric authentication technique. Some of the biometric based techniques include facial, palm prints, retinal and iris scans, hand geometry, signature capture and vocal features. Identity verification or authentication of any person is usually done using things like what you know (that is, password, PIN, etc), what you have (that is, a key, a passport, an identity card or similar possessions) or what you are (biometric feature).

What you know such as password could be forgotten or misused as possession of a token or password disclosure to the wrong individual can lead to security violations. With the new technological advances, biometrics, a new identity verification method, is becoming more popular. Instead of having to rely on what a person has or knows, physiological characteristics or personal behaviour traits, known as biometrics are getting popular in this present era.

A pattern recognition system which recognizes a user by determining the authenticity of a specific anatomical or behavioural characteristic possessed by the user is defined as biometric system (Pfeuffer et al., 2019) in order to develop a biometric system, it is very important to consider three steps namely user enrolment, user verification and biometric system testing. User enrolment: at this step of biometric system development the biometric feature of the user is captured and stored in the system as a template for reference purposes.

User verification/identification: at this stage of development, the template stored during enrolment stage is used for matching when a particular user is to be authenticated/identified. Biometric system can be used in two different modes: verification/authentication and identification Verification/authentication is a 1:1 matching in which the system performs a one-to-one comparison of a captured biometric with a specific template stored in a biometric database in order to verify the individual is the person he/she claim to be. Identification on the other hand is a 1: n search in which the system performs a one-to-many comparison against a biometric database in attempt to establish the identity of an unknown individual. The system succeeds in identifying the individual if the comparison of the biometric sample to a template in the database falls within a previously set threshold (Gangopadhyay, Chatterjee, & Das, 2019). It was established that identification mode can be used for either positive (that is, the user does not have to provide any information about the template to be used) or for negative recognition (where the system establishes whether the person is who he/she (implicitly or explicitly) denies to be).

Testing: This process may use a smart card, username or ID number (e.g. PIN) to indicate which template should be used for comparison. A typical biometric system described by (Gowda, Kumar, & Imran, 2019) is shown in Figure 1. the components in the system include sensor, pre-processing, feature extractor, matcher and application device. Sensor: Sensor serves as an interface between the real world and the system; it is used to acquire all the necessary data. Most of the times it is an image acquisition system - such as camera, fingerprint scanner etc. - but it can change according to the characteristics desired.
Pre-processing: pre-processing stage removes all the artefacts, such as background noise, from the image captured by the sensor in order to enhance the input. Some image processing technique such as normalization, histogram equalization or histogram specification among others are employed here. For instance, in fingerprint,

pre-processing involves the conversion of the fingerprint image into a usable and comparable format that does not require lots of storage space.

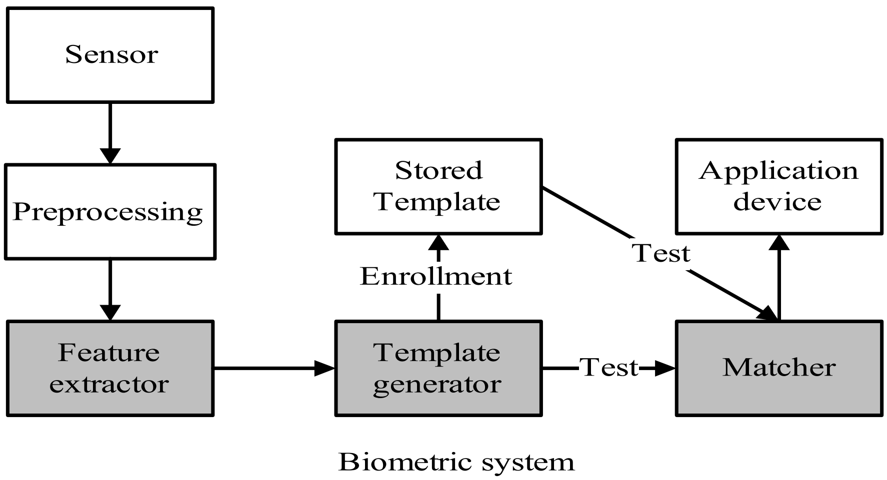

Figure 1: A typical Biometric System (Jain, Ross, & Arun, 2008).

Feature Extraction: This step is an important step as the correct features need to be extracted in the optimal way. A vector of numbers or an image with particular properties is used to create a template. A template is a synthesis of the relevant characteristics extracted from the source. Elements of the biometric measurement that are not used in the comparison algorithm are discarded in the template to reduce the file size and to protect the identity of the enrolee. Different technology demands different extractions to be made. For example, in fingerprint, an algorithm may search the image and eliminate one of two adjacent minutiae, as minutiae are very rarely adjacent.

During enrolment, the template is simply stored somewhere (on a card or within a database or both). If a matching phase is being performed, the obtained template is passed to a matcher that compares it with other existing templates, estimating the distance between them using any algorithm (such as Hamming distance). The matching program analyses the template with the input. This will then be output for any specified use or purpose.

2.14 Forms of Face Recognition

Yang et al. (2008) has generalized various face detection approaches into four main categories including:

- **Knowledge-based methods**: Prior human knowledge about face images are applied to generate a set of rules. These rules are used to determine face region candidates according to a coarse to fine multi-resolution hierarchy. Often, the relationships of facial features (e.g., eyes, mouth, nose and face contour etc.) are encoded in these rules.

- **Feature invariant approaches**: Pixel level features, such as edges, texture and skin-colour are detected from face image first. Face regions are determined based on the relationships and characteristics (such as geometric shape and morphologic relationships) of these detected features.
- **Template Matching methods**: A face template is created manually or through statistical approaches based on a training face image database. Faces are then determined using the correlation score between the face template and face region candidates.
- **Appearance based methods**: A face/non-face classification model is trained using statistical algorithms to capture characteristics of the holistic face appearance. Face regions are determined by the output score of the classification model.

However, the Knowledge-based face detection approach was contributed by (Li, Zeng, Li, Yu, 2019), where rules were considered in three different levels. On the highest level, general descriptions of a face such as shape and average intensity values were applied to scan for possible face candidates from the image. The decision is then refined by using lower level rules to capture details of facial features such as eyes and mouth. Based on the rule-based approach, Kotropoulos and Pitas proposed a face detection algorithm and it was evaluated on the M2VTS dataset which contains video sequences of 37 different subjects. 86.5% detection rate was reported under a well-controlled experiment condition.

The face detection method by (Sagar & Narasimha, 2019) was among the earliest attempts of the Template Matching approach. Several templates, including eyes, mouth, nose and face contour, were constructed using line-segments. In the detection stage, edge detection was first applied to the face candidate. Correlation between the resulting edge map and the constructed face template was then measured to decide whether the candidate is a face. As an extension to the basic template matching strategy, the focus of attention is applied in the detection process where large-scale templates (such as face shape contour) are evaluated first. Sub-templates matching is applied to locate regional facial features such as eyes and mouth. Other extensions to the template matching approach include using Silhouettes and Principle Component Analysis (PCA).

The Knowledge-based, Feature invariant and Template Matching approaches have been widely applied in the earliest work of face detection due to their simplicity in implementation and training. These approaches are also proven to be very effective to well controlled image dataset, but their performance degrades dramatically under challenging conditions such as complicated background, complicated illumination condition, low resolution, face pose variation and multiple faces. On the other hand, the Appearance based face detection approach is much more robust to these complications with a trade-off of longer time in training the model for classification. In recent years, the rapid advancement in digital data storage and computation processors have made the Appearance based approaches feasible to many real-life applications and thus it has become extremely more popular.

2.3 Convolution Neural Network

Convolution neural network is one of the deep learning algorithms used in the computer vision applications and object classification accuracy (such as object detection and recognition). CNN is particularly deployed for real-time processing, and parallel or heavy computations (Benjdira, Khursheed, Koubaa, Ammar, & Ouni, 2019). These networks utilize a deep architecture consisting of convolutional layers, pooling layers, fully connected layers, and non-linear activation functions. This kind of structure converts the input data into a deep low-dimensional feature depiction for the purpose of facilitating the task of classification by the fully connected layers (Cameron et al., 2019).

Aside being effective feature depiction approach, CNNs have been extensively applied in the field of computer vision. Following the success story of AlexNet, several novel CNNs have evolved such as ResNet, VGG, DenseNet, etc. In fact, these have offered high performance in all computer vision tasks. The main module of CNN model is the convolution layer that is used extracts features from high-dimensional structural data in more efficient by a collection of convolution kernels (Zhu, Xu, Xu, & Chen, 2018). Whenever considering multi-channel inputs (such as color images), the convolution kernels combines the various channels by adding up the convolution results to produce one single output channel per kernel.

The foremost and notable CNN-based object detector is the Regional CNN (R-CNN), which was discovered to be more efficient when compared to conventional algorithms used in recognition or detection with a 30% improvement for the PASCAL VOC 2012 image dataset (Girshick, Donahue, Darrell, & Malik, 2014). After two years of R-CNN, several other associated algorithms continue to enhance object recognition or detection including Faster R-CNN (Benjdira et al., 2019), which offered a speedup of 250 times more over the existing R-CNN. In the case of the Faster R-CNN, a two phase algorithm is involved because of the two distinct CNNs utilized for the detection of object. The first phase is responsible for finding region proposals in order to identify potential regions of position of an object. And, the second phase is used to classify the objects as well as refine the proposals. Another study by Ramachandran et al. (2012) highlighted the problems of object detection with CNNs based approaches; in particular, the pre-CNN methods showed that the appearance of objects differs from a plethora of factors to include: lighting, orientation, the size of the objects, and occlusion (Ramachandran et al., 2012; Cameron et al., 2019).

CNN emulates the features of biological networks (Agarwal, Jain, Regunathan, & Kumar, 2019). Analogous to AI-based approaches, the typical workflow of the CNN takes an input, operates on the input by means of activation functions, and the output is produced. It is a top choice for anything image classification, natural language processing, and image recognition, and image analysis. The name CNN was coined from the convolutional movement performed on the images during one of the above mentioned operations (Bernal et al., 2018). Its basic architecture consists of an input layer, hidden layer(s), and an output layer. On a broader scope, the layers in a CNN consist of multiple convolutional layers, pooling layers, and the fully connected layers. A brief description of the functions of each layer is given as follows (Albawi & Mohammed, 2017):

i. **The convolutional layers:** As the name implies, this layer convolves the input and passes its outcome to the next layer.
ii. **The pooling layers:** The pooling layer minimizes the dimensions of the data (reduces the size of the image) by merging the results of a cluster of neurons from one layer into a single unit in the next layer. This operation can be done in two ways; max pooling or average pooling. Max pooling uses the maximum value obtained from the neuron clusters, while average pooling uses the average value obtained from the neuron clusters.
iii. **The fully connected layers:** This layer connects every unit in a layer to every unit in another layer. Its function is similar to that of the Multi-Layer Perceptron (MLP) neural network.

Other important functionalities in a CNN include:

i. **The receptive field:** In general, each neuron in the neural network receives input from some points of location in the previous layer. This point of input location is called the receptive field. The neurons in a fully connected layer receive input from all the neurons in the previous layer. In a convolutional layer, inputs are received from a portion of the previous layer.
ii. **The weight:** This represents the influence or the strength of the connection between nodes or neurons in the network.
iii. **Bias:** The bias is an input node which produces a constant value 1. It allows the behaviour of the layer to be controlled.
iv. **Filter:** Filters are used for the detection of features on an image.
v. **Stride:** The stride represents the number of pixels moved by the filter. Stride can help in reducing the image size.
vi. **Padding:** The size of images often shrinks with the application of a convolutional operation. This can lead to information loss. To curtail this issue, the image can be padded with an additional boarder. This can be done by adding a pixel to the edges of the image. This is known as padding.
vii. **Parameter sharing:** Parameter sharing is performed to control free parameters. This could be implemented by using the same weight and bias for every neuron.
viii. **Softmax layer/classifier:** This layer classifies the input passed to it from the previous layer into several categories.
ix. **Activation Functions:** An activation function is applied to the output obtained after convolving over an image with a filter. This results in the generation of activations. The various type of activation functions used by previous researchers are the sigmoid, the ReLU (Xie, Fuyong , Xiangfei , Hai , & Lin , 2015), and the Tanh function.

2.4 Related works

To improve on the performance of recognition and authentication system of CNN, the layering of the convolutional and sampling layers collapsed into a single layer was proposed by (Wang, Huang, Su, & Li, 2018). The concept relied on the optimizing the parameters of face data through pre-training with fully connected layer and softax classification layers. However, there is need for large-scale images during training phase. Sunderhauf, McCool, Upcroft, & Perez, (2014) shows a supervised plant classification gadget that makes use of features from CNN. The special components of the plants which includes branch, stem, leaf, leaf scan, fruit and flower has been referred to as content categories. The methodology includes the use of features which are pre trained. Feature extraction is done by using examining one of the initial absolutely connected layers (layer 17, layer 19). Each category uses separate classifier to train the dataset. The class result was getting by averaging output of all image capabilities that offers probability distribution for each image. The experimental result shows that layer19 accomplished higher for plant classification.

Activation functions have a large influence on the training performance of deep CNN. Four major activation functions were utilized in the work of (Doorn, 2014) including max-pooling, dropout, maxout activation function, which are used to fine tune the output of the neurons and dealing with diminishing gradient flow to the lower layers of the neural networks, and sparsity. The target is to minimize the redundancy in dataset used for recognition system.

Kibria and Hasan (2017) looks at a study on research that examine four strategies consisting of bag of words (BOW), histogram of oriented gradients (HOG), extractor with SVM classifier, convolution neural network (CNN) and pre-trained CNN with SVM for image classification to find the accuracy to perceive knives from image dataset. The technique involves the following steps, the first step offers with extracting features from image. In the second step, a classifier is trained with these features to recognize after which classify them effectively after which, the classifier predict image category. The classifier is then examined with a test set to view its accuracy. All those algorithms were tested in two ways. In the first case the dataset is divided into two equal set as training and testing set. The other case, the sets are reversed and the early is used for testing while the later for training. The end result showed that deep learning CNN (pre-trained, untrained) based strategies performed best in accuracy. It was noted that the use of pre-trained AlexNet together with SVM performed best amongst all.

The quest to improve the face quality of video streams and processing speed of face recognition system motivated the work of (Rehman et al., 2019). The approach used frame selection of key-frames extraction engine and graphic processing unit (GPU) acceleration, which is used to extract key-frames of high quality faces correctly and speedily (Zou et al., 2019). The outcomes improve the face recognition accuracy of the new CNN procedure.

3. PROPOSED AUTHENTICATION SOLUTION

The major activities to be carried out in development of candidates face authentication scheme is broadly broken down into four as shown in Figure 2.

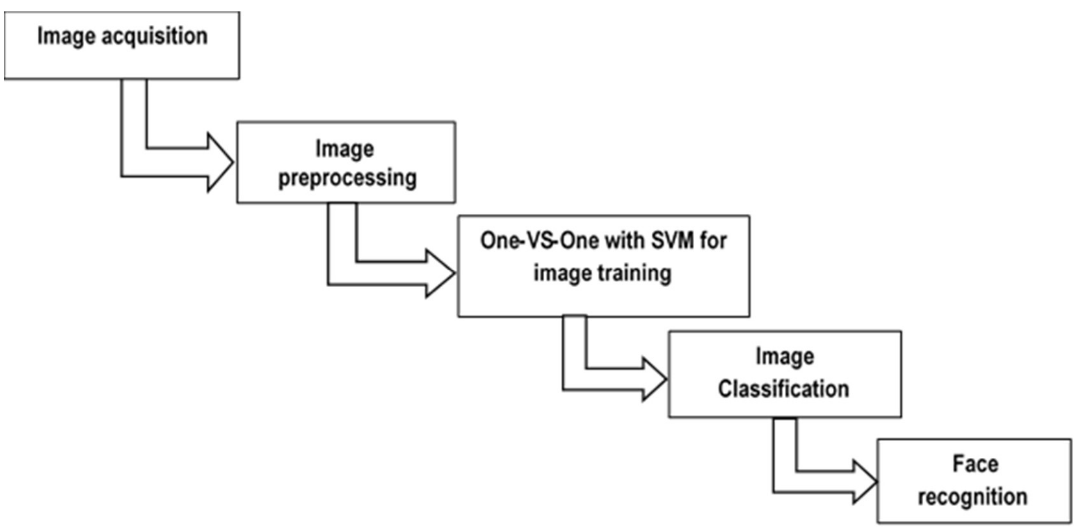

Figure 2: Proposed Research Design

From Figure 2, the image acquisition is concerned with the image capturing of students using standard devices and camera. The captured images are in raw format with several noise and redundancy. The pre-processing is to be carried out on the raw images captured using diverse techniques in order to remove the noise and redundancy for the purpose of making them fit for the next level of face recognition activities. The pre-processed image is converted to the binary form and passed through the One-VS-One approach with SVM in order to properly train the images by generating the best features for the recognition stage of the CNN.
Feature Extraction using CNN: The first and most important step deals with extracting features from all images. This is performed by creating a data store of image object which holds images along with its category labels. The labels have been assigned by taking the names of the folders that contain the image files. The feature

extraction technique used is CNN, with its ease to train a database, we used a pre trained CNN as a feature extractor. The initial layers of the network are responsible for capture basic image features (edges and blobs). Only few of the starting layers with in network (CNN) deals with feature extraction. The input image receives a response from each layer produced by CNN. To capture advanced features the early features are combined and processed by deeper layers of network. In ResNet50, the training features will be extracted using layer "fc1000".

The Image Classification activity is to be performed in the convolutional layer from the pre-trained model by setting the filter parameters, max pooling layer and reshaping of convolutional layer of the CNN. The appropriate features in the pre-trained images generated are to be passed on to the fully connected/dense layer to be used for the face recognition block.

Here are the Image Classification Steps

Step 1: Input image from one of the category folder.

Step2: Load Database: Load images use a function Image: Data Store which operates on image location to hold images and labels that are associated with each image category. An image data store allow us to store large image data, it also divide the data into 70% training and 30% test data. As each folder has different number of images per category. Using the function Count Each label determines the smallest number of images in category in order to get exactly the same number of images in each category. Load pre-trained ResNet50 Network using function resnet50.

Step 3: Image Pre-processing: The CNN model processes both training Set and test Set, where these datasets are divided into training and validation data. Thirty percent (30%) of images are randomly chosen for training and the remaining seventy percent (70%) as validating data. Image Pre-processing for CNN depending on the network used, it is performed by, Resize the image according to the network (224-by-224) and converting grayscale images into RGB images using the function augmented Image Data store.

Step 4: Extract Training Features Using CNN using function Activation with feature layer fc1000.

Step 5: Training of a multiclass SVM Classifier is done using CNN Features.

Step 6: Evaluation of the Classifier is performed by extracting image features from test dataset then these features are passed to the classifier for measuring the accuracy of the trained classifier.

The face recognition is responsible for the evaluating the performance of the CNN built for the detection of the students images during the UTME. The features selected from the original image are used to train and evaluate the face recognition system, which serves as the outcome of the proposed system.

The algorithm adaption aims to invent a new approach specifically for multi-label classification. The problem transformation focuses on transforming a multi-label classification task to a set of binary classification tasks. This approach is very popular since it can employ any classification techniques that are suitable for any domains resulting in high prediction accuracy.

3.1 Algorithm of the Proposed Model

The algorithm of the proposed model involves the steps as follows:

i. Capturing of the facial images of individual student using a digital camera
ii. Creation of database of the captured image using MySQL. Each individual has 4 images of

samples of different facial expressions.

iii. The image database will be trained using the CNN technique and the features will be saved for recognition purposes.

iv. A novel face recognition algorithm will be created based on modified CNN.

v. The modified face recognition algorithm will be used to authenticate each candidate that has to sit for the computer based test or examination.

vi. The modified face recognition algorithm will be compared with that of existing algorithms to evaluate its strength.

3.2 Performance Evaluation Parameters

The evaluation of the proposed face recognition and detection system for UTME candidates' authentication system is to adopt the parameters suggested by (Ma , et al., 2018): These include:

i. Detection accuracy: Detection accuracy or classification accuracy measures the proportion of all correct classification (for both normal and abnormal cells) in the dataset is given by Equation 1.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \qquad (1)$$

ii. Sensitivity: Sensitivity calculates the proportion of True Negative (actual positive or normal cells) that is correctly classified. The proposed algorithm is expected to have a higher degree of sensitivity as compared to existing algorithms is given by Equation 2:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (2)$$

iii. Specificity: Specificity computes the proportion of True Positive (actual negatives or abnormal cells) that are correctly classified. The proposed algorithm is expected to have a higher degree of specificity as compared to existing algorithms is represented by Equation 3:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \qquad (3)$$

Where,
TN is true negative rate,
TP is true positive rate,
FN is the false negative rate, and
FP is the false positive rate.

4. CONCLUSION

The paper considered face recognition as several technologies and scanning techniques to enhance face identification solutions. It was discovered that 2D face recognition through images captured by a standard camera is simplest and less expensive as against other methods. However, there are certain technical challenges such as system cope badly with variations in face recognition and lighting conditions are problematic and giving rise to lower accuracy rates. The proposed technique needs to be implemented in order to ascertain their effectiveness in accurately recognising faces of candidates sitting for UTME in Nigeria.

REFERENCES

1. Agarwal, R., Jain, R., Regunathan, R., & Kumar, C. S. P. (2019). Face Recognition Technique. Proceedings of the 2nd International Conference on Data Engineering and Communication Technology, Advances in Intelligent Systems and Computing, 828, 525–533. https://doi.org/10.1007/978-981-13-1610-4
2. Albawi, S., & Mohammed, T. A. (2017). Understanding of a Convolutional Neural Network.
3. Becherer, N., Pecarina, J., & Hopkinson, K. (2019). Improving optimization of convolutional neural networks through parameter fine-tuning. Neural Computing and Applications, 31(8), 3469–3479. https://doi.org/10.1007/s00521-017-3285-0
4. Benjdira, B., Khursheed, T., Koubaa, A., Ammar, A., & Ouni, K. (2019). Car detection using unmanned aerial vehicles: Comparison between faster r-cnn and yolov3. In 2019 1st International Conference on Unmanned Vehicle Systems-Oman, 1-6. IEEE
5. Bernal, J., Kushibar, K., Asfaw, D. S., Valverde, S., Oliver, A., Martí, R., & Lladó, X. (2018). Arti fi cial Intelligence In Medicine Deep convolutional neural networks for brain image analysis on magnetic resonance imaging : a review. Artificial Intelligence In Medicine, 1-18. https://doi.org/10.1016/j.artmed.2018.08.008
6. Cameron, J. A. D., Savoie, P., Kaye, M. E., & Scheme, E. J. (2019). Design considerations for the processing system of a CNN-based automated surveillance system. 136, 105–114. https://doi.org/10.1016/j.eswa.2019.06.037
7. Gangopadhyay, I., Chatterjee, A., & Das, I. (2019). Face Detection and Expression Recognition Using Haar Cascade Classifier and Fisherface Algorithm. Recent Trends in Signal and Image Processing, Advances in Intelligent Systems and Computing, 922, 1–11. https://doi.org/10.1007/978-981-13-6783-0
8. Girshick, R., Donahue, J., Darrell, T. & Malik, J. (2014). Rich Feature hierarchies for accurate object detection and semantic segmentation. In proceeding of the IEEE conference on computer vision and pattern recognition (pp. 580-587).
9. Gowda, H. S., Kumar, G. H., & Imran, M. (2019). Multimodal Biometric Recognition System Based on Nonparametric Classifiers. In Data Analytics and Learning, 269-278.
10. Hwang, K., & Sung, W. (2013). Fixed-Point Feedforward Deep Neural Network Design Using Weights + 1, 0, and -1.
11. Jain, A. K., Ross, A. & Arun, F. (2008). Introduction to Biometrics. In A. K. Jain, Flymn, & A. Ross, A Handbook of Biometrics (pp. 1-22). Springer.
12. Jones, M.J. (2009) Face Recognition: Where We Are and Where to Go from Here. IEE Journal Trans. Electronic and Information System, 129, 770–777.
13. Kibria, S. B. & Hasan, M. S. (2017, December). An analysis of Feature extraction and Classification Algorithms for Dangerous Object Detection. In 2017 2nd International Conference on Electrical & Electronic Engineering (ICEEE), 1-4. IEEE.

14. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521 (7553), 436– 444. DOI: 10.1038/nature14539
15. Li, Z. Z., Zeng, Q. H., Li, X. D., & Yu, Y. (2019, July). Face Detection Technology Based on Combining Skin Color Model with Improved Adaboost Algorithm. In 2019 IEEE 4th International Conference on Signal and Image Processing, 381-384.
16. Ma, B., Xia, Y. (2018). "Autonomous Deep Learning. A Genetic DCNN Designer for Image Classification
17. Mau, S.; Dadgostar, F., & Lovell, B.C. (2012) Gaussian Probabilistic Confidence Score for Biometric Applications. In Proceedings of the 2012 International Conference Digital Image Computing Techniques and Applications (DICTA), Fremantle, WA, Australia.
18. Pfeuffer, K., Geiger, M. J., Prange, S., Mecke, L., Buschek, D., & Alt, F. (2019). Behavioural biometrics in vr: Identifying people from body motion and relations in virtual reality. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1-12.
19. Ramachandran, U., Hong, K., Iftode, L., Jain, R., Kumar, R., Rothermel, K., & Sivakumar, R. (2012). Large-scale situation awareness with camera networks and multimodal sensing. Proceedings of the IEEE, 100(4), 878–892.
20. Rehman, B., Ong, W. H., Tan, A. C. H., & Ngo, T. D. (2019). Face detection and tracking using hybrid margin-based ROI techniques. The Visual Computer, 1-15.
21. Sagar, D., & Narasimha, M. K. R. (2019). Development and Simulation Analysis of a Robust Face Recognition Based Smart Locking System. Innovations in Electronics and Communication Engineering, Lecture Notes in Networks and Systems, 33, 3–14. https://doi.org/10.1007/978-981-10-8204-7
22. Soutar, C., Roberge, D., Stoianov, A., Rene, G., & Kumar, B. V. (1999). Biometric Encryption. In R. K.Nichols, & McGraw-Hill, ICSA Guide to Cryptography (p. Chapter 22). USA.
23. Sunderhauf, N., McCool, C., Upcroft, B. & Perez, T. (2014). Fine-Grained Plant Classification Using Convolutional Neural Networks for Feature Extraction. In CLEF (Working Notes) 756-762.
24. Sze, V., Chen, Y. H., Yang, T. J., & Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. Proceedings of the IEEE, 105(12), 2295– 2329.
25. Wang, Y., Huang, J., Su, C., & Li, H. (2018). Furnace thermal efficiency modeling using an improved convolution neural network based on parameter-adaptive mnemonic enhancement optimization. Applied Thermal Engineering. https://doi.org/10.1016/j.applthermaleng.2018.12.048
26. Xie, Y., Fuyong, X., Xiangfei, K., Hai, S., & Lin, Y. (2015). Beyond classification: structured regression for robust cell detection using convolutional neural network. In International Conference on Medical Image Computing and Computer-Assisted Intervention, 358-365. Springer, Cham.
27. Yan, S., Shan, S., Chen, X. and Gao, W. (2008). Locally assembled binary (lab) feature with feature-centric cascade for fast and accurate face detection. Proceedings of the Conference on Computer Vision and Pattern Recognition, 1–7.
28. Zhao, Y., Lang, W., & Li, B. (2019). Performance analysis of neural network with improved weight training process. 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference, 1760–1764.
29. Zhu, X., Xu, Y., Xu, H., & Chen, C. (2019). Quaternion Convolutional Neural Networks. ECCV 2018, 1, 1–17.
30. Zou, L., Yu, S., Meng, T., Zhang, Z., Liang, X., & Xie, Y. (2019). A Technical Review of Convolutional Neural Network-Based Mammographic Breast Cancer Diagnosis. Computational and Mathematical Methods in Medicine, 1–16. https://doi.org/https://doi.org/10.1155/2019/6509357