Direct Research Journal of Engineering andInformation Technology
Vol. 11 (3), Pp.29-37, March 2023
ISSN 2354-4155
DOI: https://doi.org/10.26765/DRJEIT19438419
Article Number: DRJEIT19438419

http://directresearchpublisher.org/drjeit/

Review paper

Development of a neural Network Face Biometrics System for Impersonation Detection in an Examination

Shaibu Suleman Prince*, Akhetuamen Sylvester O. and Olaniyan Julius

Department of Computer Science, School of Information and Communication Technology, Auchi Polytechnic, Auchi, Edo State, Nigeria.
*Corresponding author E-mail: princeshaibu27@auchipoly.edu.ng

Received 6 February 2023; Accepted 9 March 2023

ABSTRACT: Impersonation is one of the worst and most unsettling types of examination malpractice; it involves a collaboration between students and the examiners, lecturers, and other students to allow an unregistered (but brilliant) student to take the exam in place of a registered one. Obviously, this has a disastrous impact on the entire educational system. The construction of a software system to identify impersonators while being examined is the main focus of the paper. The system is built with Deep Convolutional Neutral Network (DCNN) biometric technology, which is based on identifying a person's face traits, and Model View Controller (MVC) image processing approach. MySQL was used as the database, and Python was used to implement it. Data was gathered through primary and secondary sources, and the face biometric dataset will be purchased online. The research resulted in the development of an impersonation detection system during examination for students at Federal Polytechnic, Auchi, in order to establish an environment conducive to free and fair examination.

Keywords: Face biometric, neutral network, convolution neutral network, model view controller

Citation: Shaibu, S. P., Akhetuamen, S. O. and Olaniyan, J. (2023). Development of a neural Network Face Biometrics System for Impersonation Detection in an Examination.Direct Res. J. Eng. Inform. Tech.Vol. 11 (3), Pp.29-37.
https://doi.org/10.26765/DRJEIT19438419

INTRODUCTION

The most common issue in the examination system is malpractice. Malpractice is difficult to detect due to the lack of a solid identity verification method for both offline and online examinations (Tejashwini, 2017). Examination malpractice has harmed the quality of graduates from many institutions in Nigeria and other parts of the world. It could range from introducing extraneous things into testing rooms to using machines to write exams instead of people (Siyaka et al., 2020). This practice jeopardizes the legitimacy of any examination involving a large number of students. To mitigate this behaviour, many schools ask students to present some sort of identification, which might be anything from a student ID card to a library card to a clearance card to a school fee payment receipt to a course registration paperwork (Vivian and Ise, 2020). Nevertheless, none of them are sufficient to determine a student's examination eligibility or authenticity because they are all easily falsifiable. The previous strategy has been ineffectual over time since students can simply impersonate one another and use false identities to get entrance to test rooms. This has had a substantial impact on the validity of examinations and graduates' fitness for post-secondary jobs (Ackerson et al., 2021). For as long as computers have existed, authenticating students to guarantee they are eligible for a specific examination has been a difficult task. Various authentication mechanisms have been developed, ranging from password to automated attendance. The best sort of authentication, however, is one that allows the user to utilize something personal. This is where biometric authentication comes into play, as these systems provide an alternative to authentication by

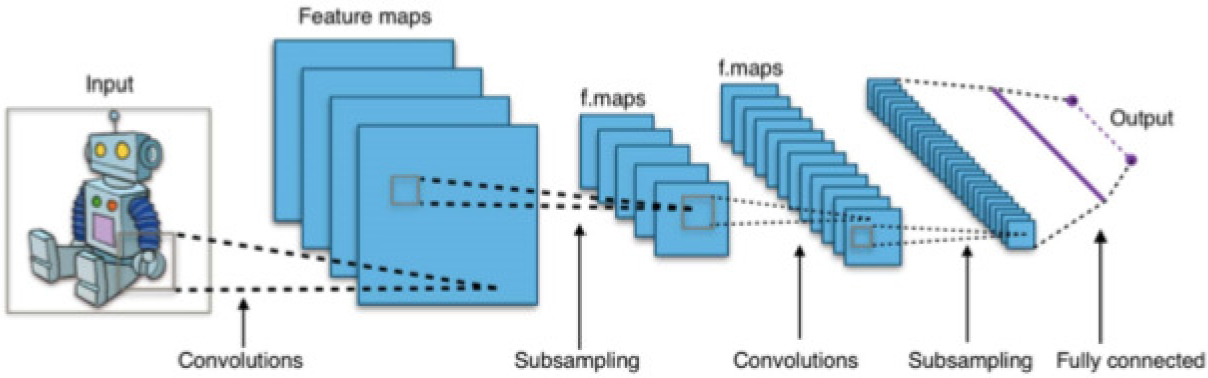
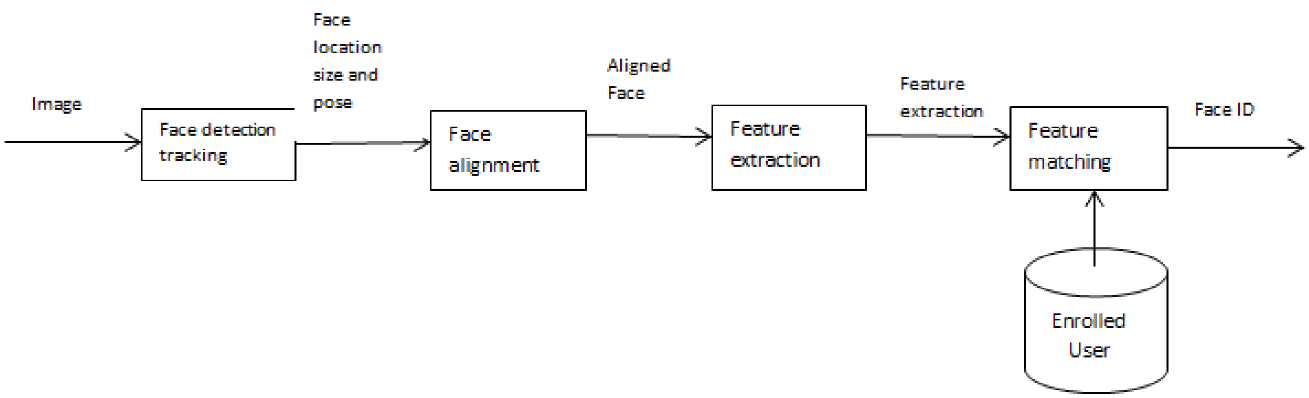


Figure 1: Typical convolutional neural network (CNN) structure (Convolutional_neural_network, 2017).



`

`**Figure 2:** Configuration of a face recognition system.

sensing physiological or behavioural characteristics and using them to identify a person (George et al., 2019). Biometrics is the science of identifying a person's personality based on an investigation of their behaviour, and physical or chemical features. Fingerprinting, face, and voice recognition are the most common biometric techniques.

A variety of procedures will be performed by the biometric system in order to recognize personality (Isinkaye et al., 2020). Face Biometric system, like other forms of identification such as passwords, email verification, selfless or photos, or fingerprint identification, employs unique mathematical and dynamic patterns that make it one of the safest and most successful. Moreover, with the help of a trained convolutional neural network, which is now state-of-the-art within the scope of this assignment, an algorithm for extracting significant features will be developed. convolutional neural network (CNN) is a deformation of multi-layer perceptron inspired by biological vision and the most simplified preprocessing operation. It is essentially a forward feedback neural network (Figure 1).

The biggest difference between convolutional neural network and multilayer perceptron is network. The first few layers are composed of a convolutional layer and a pooled layer alternately cascaded to simulate a simple cascade of cells and complex cells for high level feature extraction in the visual cortex (Sangwhan, 2017) (Figure 1).

Figure 2 depicts four modules of a face recognition system: detection, alignment, feature extraction, and matching, where localization and normalisation (face detection and alignment) are processing steps before face recognition (facial feature extraction and matching). Face detection separates the areas of the face from the background (ThaiHoang, 2011).

Facial features, such as eyes, nose, mouth, ears and facial outline, are positioned; based on the position points, the input face image is normalized with respect to geometrical properties, such as size and pose, using geometrical transforms or mutating. The face is usually further normalized with respect to photometrical features such illumination and gray scale. After a face is normalized geometrically and photo metrically, feature extraction is carried out to provide reliable information that is useful for differentiating between faces of different individuals and stable with respect to the geometrical and photometrical variations. For face matching, the extracted feature from the picture of the input face is matched against those of enrolled faces in the database; it outputs the identity of the face when a match is found with sufficient confidence or indicates an unknown face otherwise.

RELATED WORKS

According to Shivam and Graceline, (2019) face recognition system performs the following operations

Face tracking

This is to detect object of face in real time and to keep tracking of the same object. Here we use the training samples images of other objects of your choice to be detected and track by training classifier. Face tracking is a part of face recognition system. Here we can use some system algorithms to pick out specific, distinctive details about a human’s face.

Face detection

This face detection process actually verifies the image is face image or not. Detection process actually works on Haar Cascade classifier. Object Detection using Haar feature- based classifiers is an effective object detection method proposed Paul Viola and Michael Jones. It is machine learning based approach where a cascade function is trained from images. It is used to detect objects in other images.

Haar cascade classifier features

Here we calculated, the first feature selected seems to focus on the property that the region of the eyes in often darker than the region of the nose and cheeks. The second feature chosen is based on the eye is darker characteristics than the bridge of the nose. However, you do not need the same window that applies to your cheeks and other places (Figure 3).

Haar cascade classifier

Face recognition system that does capturing the image of face feature detection, extraction, storing and matching. But the difficulty occurs to lay the transmission lines in the places where the topography is bad. The authors proposed a system based on real-time face recognition that is reliable, secure and fast, and requires improvement in different lighting conditions (Figure 3).

According to Yang and Sangwhan, (2020), face recognition can be traced back to the sixties and seventies of the last century, and after decades of twists and turns of development has matured. The traditional face detection method relies mainly on the structural features of the face and the color characteristics of the face. Some traditional face recognition algorithms identify facial features by extracting landmarks, or features, from an image of the subject's face (Figure 4).

Face recognition algorithm may analyze the relative position, size, and/or shape of the eyes, nose, cheekbones, and jaw. These features are then used to search for other images with matching features. These kinds of algorithms can be complicated, require lots of compute power, hence could be slow in performance. And they can also be inaccurate when the faces show clear emotional expressions, since the size and position of the landmarks can be altered significantly in such circumstance.

Face detection and face tracking

This article Robust Real-time Object Detection (Paul, 2020) is the most frequently cited article in a series of articles by Viola that makes face detection truly workable. We can learn about several face detection methods and algorithm from this publication. The article Fast rotation invariant multi-view face detection based on real Adaboost (Hach, 2019), for the first time real adaboost applied to object detection, and proposed a more mature and practical multi-face detection framework, the nest structure mentioned on the cascade structure improvements also have good results. The article Tracking in Low Frame Rate Video: A Cascade Particle Filter with Discriminative Observers of Different Life Spans (Yuan, 2008)is a good combination of face detection model and tracking, offline model and online model, and obtained the CVPR 2007 Best Student Paper. The above three papers discussed about the face detection and face tracking problems. According to the research result in these papers, we can make real time face detection systems. The main purpose is to find the position and size of each face in the image or video, but for tracking, it is also necessary to determine the correspondence between different faces in the frame.

Face positioning and alignment

Earlier localization of facial feature points focused on two or three key points, such as locating the center of the eyeball and the center of the mouth, but later introduced more points and added mutual restraint to improve the accuracy and stability of positioning Sex. The article Active Shape Models-Their Training and Application (Cootes, 1995) is a model of dozens of facial feature points and texture and positional relationship constraints considered together for calculation. Although ASM has more articles to improve, it is worth mentioning that the AMM model, but also another I mportant idea is to

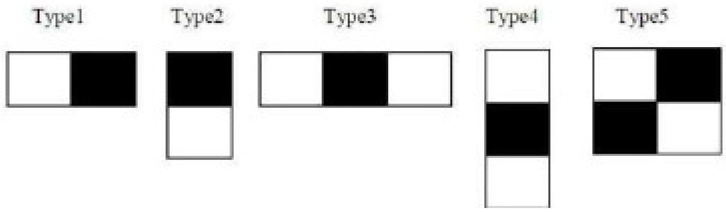


Figure 3: Haar Cascade Classifier (Shivam and Graceline, 2019).

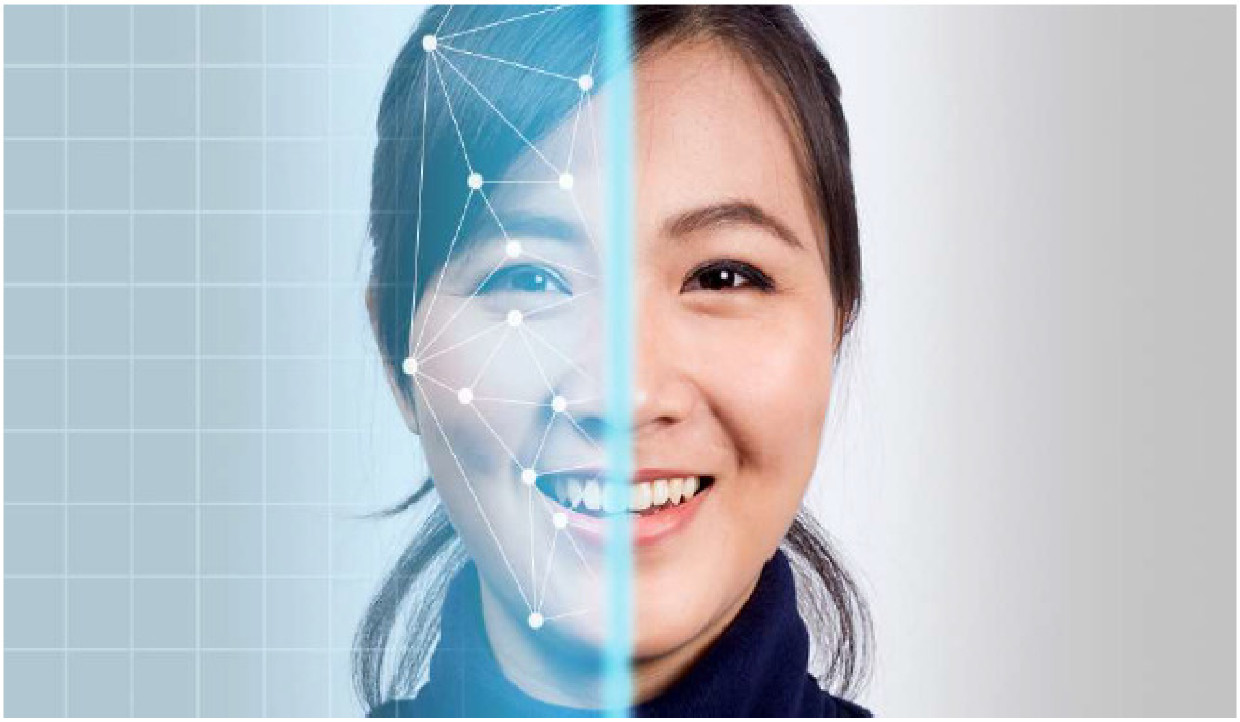


Figure 4: Abstract humane face into features (cookbook.fortinet.com accessed 2018).

to improve the original article based on the edge of the texture model. The regression-based approach presented in the paper Boosted Regression Active Shape Models (Cootes, 2020) is better than the one based on the categorical apparent model. The article Face Alignment by Explicit Shape Regression (Cao, 2012)is another aspect of ASM improvement and an improvement on the shape model itself. Is based on the linear combination of training samples to constrain the shape, the effect of alignment is currently seen the best.

The purpose of the facial feature point positioning is to further determine facial feature points (eyes, mouth center points, eyes, mouth contour points, organ contour points, etc.) on the basis of the face area detected by the face detection / tracking, s position. These 3 articles show the methods for face positioning and face alignment. The basic idea of locating the face feature points is to combine the texture features of the face locals and the position constraints of the organ feature points.

THE PROPOSED SYSTEM

The trainable Deep Learning system for facial identification and recognition created in this study is made up of a number of modules, each of which represents a different processing stage. Each module has trainable parameters that may be altered, much like the weights of linear classification models. The entire training process of the system involves adjusting all of the parameters in each example to move the system's output closer to the desired output. These modules are arranged in a series of layers to create the deep learning model. A Deep Learning architecture can be thought of as a multilayer network of basic units, comparable to linear classifiers, coupled with trainable weights in its most basic form. A multilayer neural network is what this is. Deep architectures have the advantage of being able to learn to represent the world hierarchically. There is no need to manually construct a feature extractor because every layer may be trained. That will be taken care of by the training. Additionally, the very first layers will extract basic properties (the presence of contours) that the succeeding layers will combine to create ever-more-abstract and complicated notions, such as the assembly of contours into patterns, pieces of objects into patterns, etc. Input picture properties can be automatically extracted using CNNs. It employs the idea of weight sharing, making it insensitive to minor visual distortions and allowing a significant decrease in the number of network parameters. The local correlations present in an image can also be strongly taken into account thanks to this weight sharing. A stack of independent processing layers makes up the CNN architecture starting with the convolution layer (CONV), which modifies the input data that has been received, then followed by the pooling layer (POOL), which enables information compression by minimizing the size of the intermediate picture usually by sub-sampling. Next to the pooling layer is the activation layer, also known as the Rectification Linear Unit (ReLU) activation function. This is the layer that applies a non-linear function to the input (ReLU), typically an element-wise operation whose purpose is to introduce non-linearity into the output of a neuron, allowing the neural network to learn more complex patterns in the face, in this case. The Fully Connected (FC) layer, a layer that resembles a perceptron.is the layer where every neuron in the previous layer of the CNN is connected to every neuron in the current layer. This means that each neuron in the current layer receives input from all the neurons in the previous layer, and produces an output that is passed to all the neurons in the next layer. The purpose of this layer is to learn a non-linear combination of the inputs it receives from the previous layer. The last layer of the designed network is called the classification layer (Softmax), which foretells the image's categorization. The classification layer, also known as the output layer, is the final layer in the network that is used for classification tasks. The neurons in this layer are responsible for producing the final output of the network, which is a set of probabilities or scores that correspond to the different classes or labels in the dataset. The number of neurons in the classification layer is usually equal to the number of classes in the dataset. In this research, the activation function used in this layer is softmax function, which maps the output of the neurons to a probability distribution over the classes. The class with the highest probability is then chosen as the final prediction of the network.

SYSTEM ARCHITECTURE

A CNN is suggested for facial identification in this work. Two convolutional layers, a fully connected layer, and a classification layer make up the proposed network. After each convolution layer is an activation layer and a maxpooling layer. After each convolution layer, the architecture includes batch normalization and dropout regularization algorithms. The dropout approach is used after the fully connected layer to simplify the calculations and improve the performance of the proposed CNN. The primary goal of this work is to use grayscale photos to identify and recognize student faces in an examination environment. As a result, the suggested CNN's input image size was fixed at 32x32x1, where 1 stands for the color of the grayscale space. The first convolution layer's kernel size and number of filters were fixed at 3x3 and 16, respectively. Since the input image is a grayscale image with fewer features that need to be learned, the number of filters was decreased. ReLU activation function was used for the activation layer since it is the most popular in CNNs and has demonstrated effectiveness when compared to alternative activation functions. By making the mapping function more adaptable and non-linear, the ReLU aids in preventing the generation of negative values in the event that the input image is destroyed and prevents saturation of the neurons. The kernel size was set to 2x2 for the max-pooling layers, and the stride is 2. The first convolution layer's kernel size was carried over into the second convolution layer, where the number of filters was increased to 32. A batch normalizing technique was used following the completion of each block of convolution, activation, and max pooling. With the certainty of network convergence and the ability to apply a high-value learning rate, batch normalization is a regularization technique that was used to speed up training. Additionally, batch normalization aids in initializing the CNN's weights, particularly since it was trained from scratch. Batch normalization also lessens internal covariant shift and gradient dependence on the scale of the parameters or their initial values. The addition of a dropout technique was proposed in order to prevent the over fitting issue because the input image does not have many characteristics that need to be learned. In order to concentrate on neurons with strong connections and improve performance, the dropout strategy was employed to remove neurons with weak connections. Dropout lowers the difficulty of computing by reducing the number of connections and neurons. To achieve excellent performance, regularization techniques, batch normalization, and dropout are only used during the training process. The network was expanded to include a fully linked layer that would integrate and summarize the learned information. To avoid the over fitting issue, the dropout technique was used after the completely linked layer. The Softmax function was utilized as the output layer to calculate the probabilities for each class. All class probabilities added together must equal one. The equation for the softmax function for a given input vector x is:

Softmax (x)i=exp(x_i)/sum(exp(x_j)) for all j . (1)

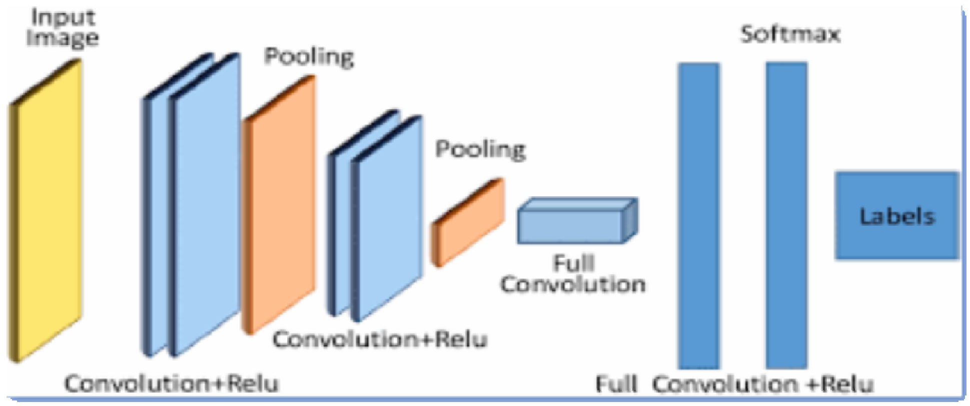


Figure 5: Generic CNN Architecture (Source: Afridi et al., 2021).

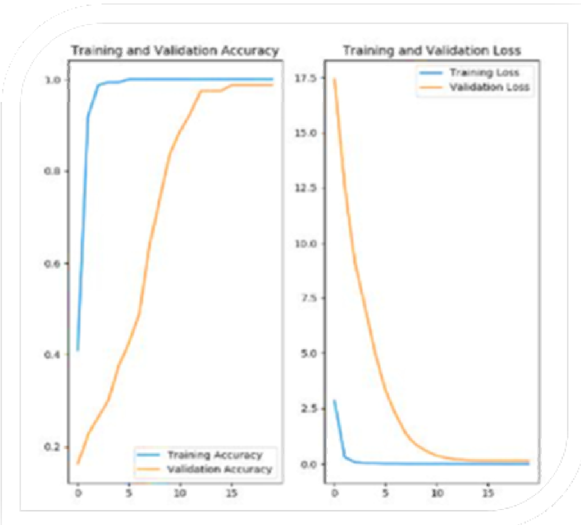
Figure 6: DigiFace-1M dataset sample.

where i is the index of the neuron in the output layer, x_i is the input to that neuron, and the denominator is the sum of the exponential of all the inputs to the neurons in the output layer. The result is a vector of the same size as x, where each element is a value between 0 and 1 that represents the probability of the corresponding class. The sum of all elements in the vector is equal to 1. This function ensures the output values are in between 0 and 1, and all sum to 1, making it a probability distribution, which allows the network to choose the class with the highest probability as the final prediction. The suggested CNN is shown in more detail in (Figure 5).

EXPERIMENTS AND RESULTS

An Idea PAd Lenovo Laptop computer with an Intel i5 processor with 8 internal cores and an Nvidia GTX 960 GPU was used for all the experiments. Using the Ten sorflow Deep Learning framework and the Open CV library, the system was fully developed in Python programming language coded with Pycharm Integrated Development Environment. Used in this research is DigiFace-1M, which is a dataset of 1 million facial images and corresponding annotations that is used for training and evaluating the designed model for facial recognition and analysis tasks. It includes a diverse range of images and annotations, such as name, age, gender, and facial landmarks, which can be used to train models for a variety of applications, such as face verification, identification, and emotion recognition. The dataset is publicly available for research purposes and can be used to train and evaluate models for a wide range of facial recognition and analysis tasks (Figure 6) presents a visual sample of the dataset as follows: The dataset is divided into training, validating and testing sets in the ratio of 6:3:1 respectively. That is, 60% of the data was used for training, 30% for testing and the remaining 10% was used for validation.

After investigating the data, categorical crossentropy was used to calculate the proposed CNN's loss function. This function is typically applied to single-label



Fig**ure 7:** Loss function optimization curves.

classification jobs, which require that each input image be a member of a single output class. It provides a sense of how inaccurate the neural network's forecast was. The equation for the categorical cross-entropy loss for a single example with true label y and predicted probability distribution p is:

loss=-sum(y*log(p))   (2)

where:
y is a one-hot encoded vector of the true label, p is the predicted probability distribution, and the sum is taken over all classes.

The predicted probability distribution is typically produced by the output layer of the neural network, which uses the softmax activation function to ensure that the predicted probabilities sum to 1 and are non-negative. The idea behind the Categorical cross-entropy is that it wants to minimize the difference between the predicted probability of the true class and 1. As the predicted probability of the true class get's closer to 1, the loss decrease, and the network learn that the current example belongs to the true class. The proposed loss function was optimized using the Adam gradient descent algorithm. The Adam algorithm updates the weights of the neural network by computing the gradient of the loss function with respect to the weights, and then updating the weights based on that gradient. The weight update can be calculated as:

m_t = beta1 * m_{t-1} + (1 - beta1) * grad……………(3)

v_t = beta2 * v_{t-1} + (1 - beta2) * grad^2……………(4)

m_t_hat = m_t / (1 - beta1^t…………………………….(5)

v_t_hat = v_t / (1 - beta2^t)……………………………(6)

w_t = w_{t-1} - learning_rate * m_t_hat / (sqrt(v_t_hat) + epsilon)………………….(7)

where:
beta1 and beta2 are the decay rates for the moving averages of the gradient and the squared gradient respectively.

They are typically set to values less than 1, such as 0.9 or 0.99, to give more weight to more recent gradients.
m_t is the moving average of the gradient, m_{t-1} is the previous moving average, and grad is the current gradient calculated with respect to the loss function.
v_t is the moving average of the squared gradient, v_{t-1} is the previous moving average, and grad^2 is the current gradient squared.
m_that and v_that are the bias-corrected moving averages.

zw_t is the updated weight, w_{t-1} is the previous weight, learning rate is a scalar hyperparameter that controls the step size of the update, and epsilon is a small constant added to the denominator to ensure numerical stability.
Adam algorithm uses both the moving average of the gradient and the moving average of the squared gradient, which allows it to adapt the learning rate on a per-parameter basis.

The algorithm adjusts the learning rate based on the historical gradient and the historical gradient squared, which makes the algorithm more robust to noisy gradients and helps the network converge faster and more efficiently. Figure 7 displays the loss optimization and accuracy curves. On the testing set, the loss function has a minimum value of 0.14. The suggested CNN achieves an inference speed of 231 frames per second, a validation accuracy of 96.7%, and training accuracy of 98.98%. The findings collected demonstrate the effectiveness of the suggested CNN for biometric face identification. We contrast the proposed strategy with state-of-the-art approaches in order to demonstrate its effectiveness. In (Table 1), the suggested method for face identification is compared to the most recent methods in terms of accuracy. We can see that the suggested method reaches the highest accuracy performance standards. Table 1 was used to establish the robustness of the proposed method with respect to face recognition and identification tasks. This is further demonstrated in (Figure 8), which is a visual representation of the proposed method accuracy as compared with other state-of-the-heart face recognition techniques.

IMPLEMENTATION

The model created in this research was implemented and tested, and the following output were captured as presented in Figure 9. For demonstration purposes, 100

Table 1: Face recognition methods comparison.

Method	Accuracy (%)
PCA	98.3
ICA	97.75
Eigenface	97.5
SIFT	97
Proposed	98.98

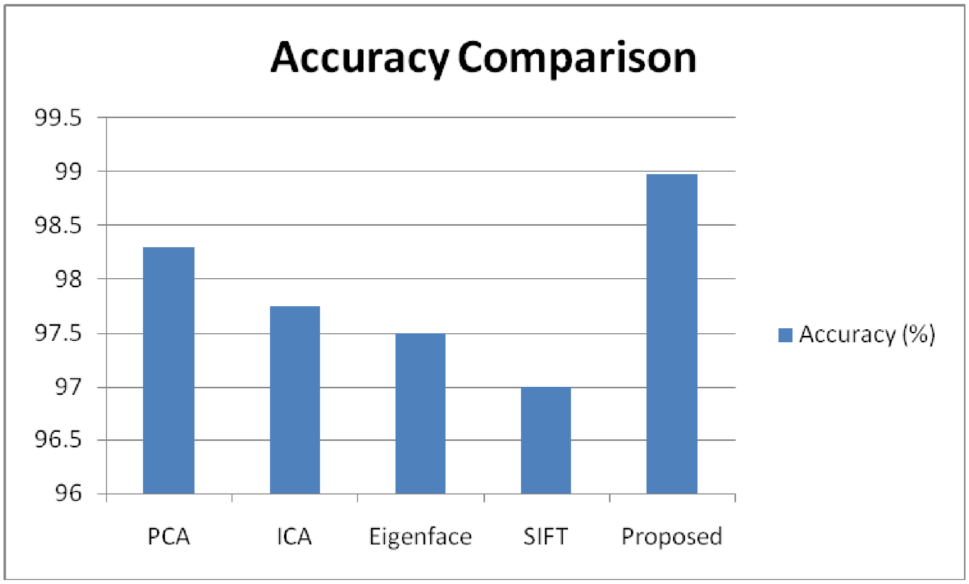


Figure 8: Face recognition methods comparison

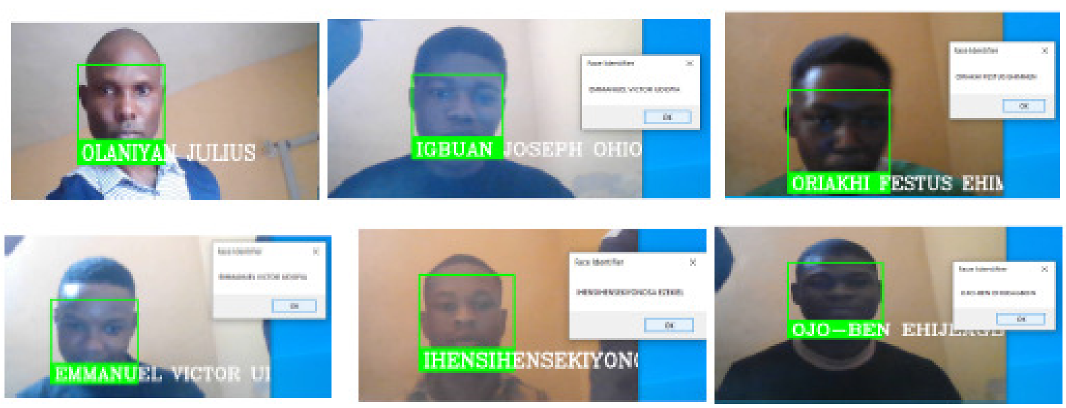


Figure 9: Recognized faces.

students passport photographs from computer science department, Auchi Polytechnic, Auchi, were stored with their respective names. Then, 6 out of the 100 who were physically available were made to stand in front of the laptop in which the designed application was installed for recognition. The system was able to perfectly recognize each of the students by labeling each webcam based identified face with the appropriate name.

CONCLUSION

One of the most crucial jobs for applications like Examination Malpractice Control System (EMCS) is face identification and recognition. Face recognition technology has the potential to be used as a tool to control examination malpractice in education. By using facial recognition to verify the identity of students during exams, educators and administrators can ensure that the person taking the exam is the same person who is registered to take it. This can help to prevent cheating and other forms of academic dishonesty. In this paper, concerted efforts have been made to design a convolutional neural network-based face identification and recognition application. The proposed CNN performs with great accuracy, which make it suitable for the task of examination malpractice control.

However, it is important to note that the use of facial recognition technology for examination control raises several ethical concerns. One of the main concerns is related to privacy, as the use of facial recognition technology in the classroom may be seen as an invasion of students' privacy. Additionally, there are concerns about the accuracy of the technology and the potential for bias, which could lead to false accusations and other forms of discrimination. Also, another important point to consider is the cost of such technology and its maintenance. Furthermore, there is a need for proper regulation and ethical guidelines to be in place for the use of facial recognition technology in education.

To sum up, while face recognition technology has the potential to be an effective tool for controlling examination malpractice, it must be approached with caution and with the consideration of ethical concerns, such as privacy and accuracy, and with the development of proper regulations and guidelines.

REFERENCES

Ackerson, J. M., Dave, R., and Seliya, N. (2021). Applications of recurrent neural network for

Afridi, T. H., Alam, A., Khan, M. N., Khan, J., and Lee, Y. K. (2021). A multimodal memes classification: A survey and open research issues. In *The Proceedings of the Third International Conference on Smart City Applications* (pp. 1451-1466).Springer, Cham.

Application," COMPUTER VISION AND IMAGE UNDERSTANDING, pp. 38-59, 1995.

biometric authentication and anomaly detection. *Information*, *12*(7), 272.

Cao Y. X., "Face alignment by Explicit Shape Regression," in 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, 2012.

Cootes D. C. a. T., Boosted Regression Active Shape Models, BMVC, 2020

Cootes H. C., "Active Shape Models-Their Training and

George, A., Mostaani, Z., Geissenbuhler, D., Nikisins, O., Anjos, A., and Marcel, S. (2019). Biometric face presentation attack detection with multi-channel convolutional neural network. *IEEE Transactions on Information Forensics and Security*, *15*, 42-55.

Hach a. S. L. Bo Wu, "Fast rotation invariant multi-view face detection based on real Adaboost," Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 79-84, 2019.

Isinkaye, F. O., Soyemi, J., and Arowosegbe, O. I. (2020). An Android-based Face Recognition System for Class Attendance and Malpractice Control. *International Journal of Computer Science and Information Security (IJCSIS)*, *180*(1), 78-83.

Paul Viola M. J., "Robust Real-time Object Detection," International Journal of Computer Vision, pp. 137-154, 2020.

Sangwhan Cha, Assistant Professor of Computer Science, harrisburgu University “face recognition system” 18 "Convolutional_neural_network," 2017. [Online]. Available: https://en. Wikipedia.org/wiki/convolutional neural networks

Shivam Singh and Graceline Jasmine, Face Recognition System, International Journal of Engineering Research and Technology (IJERT), ISSN: 2278-0181 Vol. 8 Issue 05, May-2019

Siyaka, H.O., Owolabi, O. and Hashim, B.I. (2020). An Examination Conduct Mechanism for Nigeria Based on Modified Convolutional Neural Networks Algorithm Authentication Scheme. Journal of Behavioural Informatics, Digital Humanities and Development Research.Vol.6 .No. 1, Pp 65-76. Available online at https://www.behaviouralinformaticsjournal.info

Tejashwini, S. G. (2017). Fraud Detection in Examination using LBP method. *International Journal of Latest Engineering and Management Research (IJLEMR). International*, *2*(4), 28-35

ThaiHoangLe, (2011). Applying Artificial Neural Networks for Face Recognition, Hindawi Publishing Corporation Advances in Artificial Neural Systems Volume 2011, Article ID 673016,

verification using deep neural network and support vector machine (SVM). *Int J. Sci. Res. Comput. Sci. Eng Inf. Technol*, *6*(4), 11-20.

Vivian, N. I., and Ise, O. A. (2020). Face recognition service model for student identity

Yang Li and Sangwhan Cha, (2020).face recognition system. harrisburgu.edu "cookbook.fortinet.com," 10 10 2018.[Online].Available: https://cookbook.fortinet.com/face-recognition-configuration-in-forticentral/.[Accessed10 10 2018].

Yuan Li, "Tracking in Low Frame Rate Video: A Cascade Particle Filter with Discriminative Observers of Different Life Spans," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1728-1740, 2008.