A Bimodal Approach for Partially Occluded Face Detection and Recognition for Crime Control in Nigeria Using Deep Learning and Machine Learning Algorithms

Odulaja Godwin

Tai Solarin University of Education, Ijagun, PMB 2118, Ijebu Ode, Nigeria

DOI: https://doi.org/10.62277/mjrd2025v6i20015

ARTICLE INFORMATION

Article History

Received: 18th April 2025
Revised: 26th May 2025
Accepted: 30th May 2025
Published: 30th June 2025

Keywords

Facial Recognition Histogram of Oriented Gradients

Multi-Task Convolutional Neural Networks (MTCNN)

Partial Occlusion; Support Vector Machine (SVM)

Visual Geometry Group Network (VGG-Net)

ABSTRACT

For the purpose of crime prevention and control, much effort has been made in literature on accurate face recognition using several approaches. However, little had been achieved on accurate identification of partially occluded faces, which is now the growing trend among criminals, as literature reveals. In this study, first, Deep Learning Multi-Task Cascaded Convolutional Neural Networks were used for face detection and face alignment, while VGG16 Convolutional Neural Network architectures were used for feature learning and classification. Secondly, and for result comparison, the Machine Learning Histogram of Orientated Gradients (HOG) with the Support Vector Machine algorithm was used for face detection as well. The feature vectors generated by the HOG descriptor were used to train Support Vector Machines (SVM), and the results were validated against given test input. The model was trained with datasets obtained from Disguised Faces in the Wild and with primary data of African facial images (occluded and non-occluded) comprising diverse occlusion patterns. Obtained results showed that the Convolutional Neural Network produces a recognition accuracy confidence level of 96% for occluded faces as opposed to the Histogram of Gradients. Convolutional Neural Network is recommended therefore for detecting and recognising partially occluded faces. For improved performance results, using larger datasets is recommended.

*Corresponding author's e-mail address: goddyseyi@gmail.com (Godwin, O)

1.0 Introduction

Crime and corruption have become multifaceted and deeply enshrined into the social fabric of Nigerian society, such that it requires more than the traditional approach to manage it. According to Adegoke (2014), in a culture like ours, crime must be taken seriously due to the mental and social problems it has caused many victims. Intelligence gathering towards crime control has become challenging due to multiple factors (Adegoke, 2020). Crime in the form of banditry, cattle rustling, kidnapping, money laundering, ritual killings, human trafficking, and terrorism had continued to menace Nigerians despite the adoption of various modern technologies such as facial recognition, surveillance, automated plate-number recognition, Geographical Information System (GIS), Close Circuit Television (CCTV), and biometrics recognition security techniques to predict criminal activities (Mohamed, 2018).

It is reported that not less than 150 people were violently killed by herdsmen and terrorists in April 2025 in Plateau and Benue states. This is despite governmental efforts and promises to stem crime (WION, 2025). Other states like Borno, Nasarawa, Gombe, Kaduna, Zamfara, and Adamawa have had pathetic and wanton destruction of lives and properties over the past 10 years, pushing survivors to become destitute, internally displaced, homeless, jobless, and poverty-stricken asylum seekers within the nation. This situation has dealt inevitable blows to Nigeria's economy and security infrastructure.

In 2023 alone, attacks from terrorist groups like Boko Haram and the Islamic State West Africa Province (ISWAP) in Adamawa, Borno, Kaduna, and Yobe states caused the death of 2,212 persons. Reports further confirmed that *Nigeria* lost a total of 11,794 individuals to *violence* in *2023* and 15,493 in 2022 to rural banditry and kidnappings (NigeriaWatch, 2023). While crime in Nigeria has reached the level of a full-blown security crisis, it has only received limited attention (Kleffman et al., 2024).

Ayegbusi (2024) reported that while kidnapping had been reported in all parts of Nigeria, countermeasures from both state and federal governments alone had failed for decades to bring the desired relief. He thereby called for collaborative efforts among all stakeholders.

Rosenje and Adeniyi (2021) attributed the prevalence of banditry and kidnapping in Nigeria to overwhelming unemployment, weak security systems, and porous borders. Worse still is the fact that senior security officers were indicted in some of these criminal cases. The weak and biased legal system of the country indirectly fosters crime.

Significantly, it is often reported that leaders of perpetrators of these nefarious activities do speak on social media to demand ransom with their faces masked. Criminal organisations like Boko Haram, Lakurawas, and the like, who have terrorised Nigerians for many years, are not excluded in this show of confidence and impunity.

In recent times, it has become common for security officers to be heard calling out to the populace for assistance in combating crime since the government's efforts like dialogue, disarmament, and lacklustre engagements had failed to address these problems. In fact, these terrorists have been emboldened to even attack military personnel, convoys, and barracks (Aladeselu, 2024).

Failure of efforts of the federal government had been attributed by Oche & Robert (2024) to lack of coordination, resource constraints, and injustices. However, to be productive, security agents need not only to know where, when, and how the terrorists operate in real time but also to be able to identify them wherever they are found, whether masked or not. The fact remains that these criminals do not always remain masked, as they come to the public for functions and for business engagements unmasked. They only mask up to hide their identity during criminal operations or on social media. Embracing technology in the form of Artificial Intelligence for occluded face recognition will be of great assistance in this regard. Oche and Robert also recommend investing in technology-driven solutions (Oche & Roberts, 2024).

In the context of this study, face recognition is a biometric recognition that consists of identifying the human face using a similar computational model. (Bachhety et al., 2018).

Advances in computing have enabled machines to accomplish facial recognition automatically. If a criminal's occluded face can be accurately identified and recognised regardless of his disguise, apprehending him later, when he least expects it, will be much easier.

Face recognition requires, in sequence, detecting the face, extracting facial features, and classifying extracted features before face recognition. Face detection implies confirming that a human face exists in a given image, as well as the location of the face. While the feature extraction phase entails extraction of peculiar human face patches from the images (Bansal, 2018), the face recognition phase is about realising whose identities the faces represent; for this to be successful, features of the target face need to be extracted first. A repository of faces is needed. For each candidate person, several images of him are captured, and his features are extracted and stored in the repository (Bansal, 2018). In many fields, such as security, education, financial industries, and others, face recognition has become the integral biometric technique for identifying and authenticating individuals. Face recognition has been used in a variety of fields, including criminal investigation, automatically detecting an object, tracking it, classifying or identifying it, and analysing its activities (Chen et al., 2018). While other biometrics, such as fingerprints and iris recognition, have featured well in some situations, their recognition is considered compromised (Martínez, 2021). Recently, face recognition has received more attention as an advanced means of crime detection.

1.1 Biometric Recognition System

Every Biometric Recognition System (BRS) must be able to achieve two major tasks: verification and identification. Biometric verification requires a one-to-one comparison of the query template of a subject with the template of the claimed identity stored in a database. On the other hand, biometric identification is a one-to-many comparison between an unknown subject and the templates stored in the database.

To foster a fast one-to-many identification process, it is a practice to divide the biometric templates kept in the database into subcategories. The system returns the identity of the unknown subject after the comparison process. The identification task is more challenging compared to the verification process. The sought-after subject must be identified first before verification follows as a means of validating the identification. Thus, identification requires more time for the one-to-many matches to complete, depending on the number of database templates to be considered during the matching process. Besides, the chances of having false acceptance are much higher at this stage (Bachhety et al., 2018).

Computer vision is the sub-domain of Artificial Intelligence (AI) that directly addresses issues of biometric identification and recognition. It utilises advanced research in machine learning and deep learning algorithms like Histogram of Orientated Gradients (HOG), Support Vector Machine (SVM), Convolutional Neural Networks (CNN), and many others. As a result, revolutionary benefits had been recorded in image & video recognition, media-based entertainment & recreation, natural language processing, image analysis & classification, recommender systems, and many more (Prabakaran and Mitra, 2018).

The architecture of CNN is patterned after the activities of neurones of the human nervous system, especially the brain. In the brain, the organisation of the visual cortex is synonymous with that of a CNN. A neurone's individual response to detected stimuli is only limited to the receptive field. As the collections of such fields overlap, the entire visual area is eventually covered (Valueva et al., 2020). The visual cortex aptly demonstrates this.

Wang et al. (2018) maintained that the most common use for CNNs is image classification. According to Kostomarove (2019), functional parallelism among all its elements gives neural networks an edge as a deep learning algorithm. As computer programs mathematically simulate this, much machine learning is achieved in many areas like natural language processing, the science of forensics, understanding human emotion and facial expressions, and criminal search and investigations.

Because of its exceptional accuracy and minimal required computation power, SVM has become many people’s preference when it comes to tasks requiring classification of objects and regression. Its focus in use is to find one hyperplane out of all available N-dimensional space (where N is the number of extracted or available features) that uniquely classifies the data points (Zhao et al., 2017).

Partial occlusion had remained a recurring challenge menacing face recognition systems, as some parts of the target image are obscured and could not be captured for training. This presents a problem for facial recognition systems, which require the entire face to be free from obscurity; otherwise, they produce incorrect classification results (Satonkar et al., 2011). According to Min et al. (2014), it is difficult to obtain satisfactory accuracy of partially occluded faces where faces are disguised to deceive the security system.

Faces can be occluded as a result of using sunglasses, masking the face for protection from communicable diseases, using nose masks, makeup techniques, wearing a beard/hair/wig, or wrapping the face with a scarf or other accessories. Detecting and recognising partially occluded faces is important given the fact that criminals disguise themselves with it to commit crimes. Concern is growing in connection with the Automated Teller Machines (ATM), where face distortion for the purpose of criminal intent has increased. It is increasingly difficult to specify how people must dress before being allowed to transact via ATM. This challenge becomes more complex when we consider miscreants' use of 3-D printed face masks, makeup techniques, and face masks.

1.2 Related Works

1. *Shanshan et al. (2016) developed a face recognition system that utilised CNN as the feature extractor and SVM for recognising faces, using* images obtained from the Casia-Webfaces database for pre-training and from the FERET database for training and testing. The CNN was trained first with ancillary data to get the updated weights and later with the target dataset to extract more hidden facial features. They used SVM as their classifier instead of CNN to recognise all the classes. With their input of facial features extracted from CNN, SVM will recognise face images more accurately. Although their model spent less training time and obtained a high recognition rate, it was not trained to identify partially occluded faces.
2. Yang et al. (2018) proposed an approach that used deep representation for partially occluded face verification. The convolutional neural network was trained to verify between the occluded and non-occluded faces for the same identity. It could learn both the shared and unique features based on multiple convolutional neural network architectures. They combined joint loss function and alternating minimisation approach to train and test the convolutional neural network. While their experimental results on the publicly available datasets (LFW 99.73%, YTF 97.30%, and CACD 99.12%) show that the proposed deep representation approach outperforms the state-of-the-art face verification techniques, they, however, did not train with data from Africa or data from non-Caucasian origins to avoid bias.
3. Chethana *et al.* (2021) presented Detection of Partially Occluded Faces Using Convolutional Neural Networks in a constrained environment. The proposed approach is experimented on the Disguised Faces in the Wild (DFW) dataset. The model shows significant performance in terms of improved recognition accuracy of 93%. They, however, did not train with data from Africa or data from non-Caucasian origins to avoid bias.
4. Similarly, Abbas (2021) proposed an occlusion-invariant technique-based face recognition system for surveillance. The colour and texture features are employed in the development of the FRS-OCC system, followed by an incremental learning algorithm (Learn++) to choose more useful features. After that, a human face is recognised using the trained stack-based autoencoder (SAE) deep learning algorithm. In comparison to existing state-of-the-art methods, the proposed methodology outperformed and obtained Sensitivity (SE) of 98.82%, Specificity (SP) of 98.49%, Accuracy (AC) of 98.76%, and Area under the Receiver Operating Curve (AUC) of 0.9995. The results show that any surveillance application can use the FRS-OCC system. They, however, did

not train with data from Africa or data from non-Caucasian origins to avoid bias.

5. In another work, Ge, Huicheng, and Jiayu (2022) proposed and applied a multiscale segmentation-based mask learning (MSML) network, consisting of a face recognition branch, an occlusion segmentation branch, and hierarchical elaborate feature masking operators to tackle occlusion problems in face recognition systems. Although their work recorded significant improvement over many others, their dataset comprised mainly non-African input.
6. Also, Kortylewski *et al.* (2020) developed an approach based on the combination of deep neural networks with the composition models for robust object classification under occlusion. The combined model recognises occluded objects, even when it has not been exposed to occluded objects during training, while at the same time maintaining high discriminative performance for non-occluded objects.
7. Nakib *et al.* (2018) Crime Scene Prediction by Detecting Threatening Objects Using Convolutional Neural Network. Crime scene prediction without human intervention can have an outstanding impact on computer vision. The study presents CNN (Convolutional Neural Network) in the use of detecting a knife, blood, and a gun from an image. Detecting these threatening objects from an image can give us a prediction of whether a crime occurred or not and from where the image was taken. The study emphasisedthe accuracy of detection so that it hardly gives us wrong alerts to ensure efficient use of the system. This model uses Rectified Linear Unit (ReLU), Convolutional Layer, Fully Connected Layer, and the dropout function of CNN to reach a result for the detection. We use TensorFlow, an open-source platform, to implement CNN to achieve our expected output. The proposed model achieves 90.2% accuracy for the tested dataset. This study, however, did not cover cases of partially occluded faces.
8. Similarly, Sreelakshmi & Sumithra (2019) proposed an effective Facial Expression Recognition model that can handle partial occlusions and pose variation using the CNN architecture MobileNet. The model achieved an accuracy of 92.5% on the occluded images. Another study on occlusion-robust face recognition is presented by Wan & Chen (2019). They present a face recognition method based on mask learning. In this paper they propose a trainable module called MaskNet for an occlusion-robust face recognition system with the datasets on the AR dataset with real occlusions. The module can be easily optimised with end-to-end training when included in existing CNN architectures. Qualitative experiments show that the MaskNet can distinguish the useful facial areas from the occluded parts well.

Vivian and Ise (2020) aimed to mitigate examination impersonation by simple face scan using a mobile phone and also to make such a model accessible and reusable for seamless integration with any kind of student identity verification project. Impersonation in the context of examination is a situation where a candidate sits in an examination for another candidate, pretending to be the real candidate. In many institutions in Nigeria, to mitigate this act, students are expected to present a means of identification before entering the examination hall. However, this approach is not sufficient to determine the eligibility of a student for an examination, as these means of identification can easily be falsified. This paper, therefore, develops a face recognition web service model for student identity verification using Deep Neural Network (DNN) and Support Vector Machine (SVM).

Surajit (2017) investigated object detection for crime scene evidence analysis using deep learning. They presented a Faster R-CNN (Region-based Convolutional Neural Network)-based real-time system, which automatically detects objects that might be found in an indoor environment. To determine its effectiveness, the study was applied to a subset of ImageNet containing 12 object classes and the Karina dataset. On average, they obtained an accuracy of 74.33%, with a mean object detection time of 0.12 s in Nvidia-TitanX.

It is in the light of the foregoing that this study developed a bimodal partially occluded faces detection and recognition system for crime prevention using Convolutional Neural Network (CNN) on one hand and Support Vector Machine (SVM) on the other hand.

2. Materials and Methods

In this study, Multi-Task Convolutional Neural Network (MTCNN) was used for face detection, face alignment, and face cropping. VGG16, with Convolutional Neural Networks (CNN) architectures, was used for learning the features before classification for the system. VGG-net was the base model to initialise the network before adding more layers for the purpose of fine-tuning the face recognition task.

Histogram of Orientated Gradients (HOG) with SVM (a feature-based machine learning algorithm) was also implemented in this study in order to compare its accuracy and effectiveness with CNN (a deep learning algorithm). The model was trained with datasets obtained from Disguised Faces in the Wild (DFW). To remove bias, primary data of facial images of Africans (occluded and non-occluded) comprising diverse occlusion patterns—sunglasses, scarves, masks, makeup techniques, hair, and other accessories used to hide originality—were also used for training.

2.1 Data Collection, Preprocessing, and Training

The primary data were collected using the rear camera of an Apple iPhone 7 Plus with 12 million pixels to improve recognition outcomes. Lossless de-noising was achieved using Wiener Filters (known to be one of the best filtering techniques, as it does not affect the image quality but reinforces the smoothness of the image taken for examination) to reduce noise without losing image features (edges, corners, and other sharp structures). The goal of de-noising is to remove the noise from the image and recover the original image.

The images were transferred to the memory of a laptop for further preprocessing. The manually collected high-quality face images were for 15 participants, and each participant has a total of 30 face images (15 non-occluded and 15 partially occluded images, respectively). Therefore, the dataset consists of 225 non-occluded images and 225 partially occluded images, giving a total of 450 face image datasets (Figure 1). The images were preprocessed by labelling each with the identity of the person in the image and resizing them to a consistent resolution, normalising pixel values, and removing noise from the face images using Wiener filters (Figure 1). Similarly, both datasets were divided into two subsets for training and testing at a ratio of 30:70, respectively.

Face detection: Because of its accuracy, robustness, and speed, a CNN-based face detection model—Multi-Task Cascaded Convolutional Neural Network (MTCCNN) selected and trained to identify face-like patterns and generate bounding boxes around detected faces.

Figure 1
Primary Dataset of African Origin

Feature extraction stage: the Convolutional Neural Network (CNN) with the VGG16 architecture model was designed to extract discriminative features from aligned face images. In the case of the histogram of orientated gradients (HOG) with SVM, HOG was used to extract the features from the images, and they were stored in the form of a feature vector matrix.

Face Alignment: To improve the performance of subsequent recognition tasks and ascertain that all faces have a consistent look, alignment of each detected face was done by normalising their orientation and scale. Because machine learning models tend to train faster on small-sized images, each image was resized to 224x224, and the colours of the images were converted to RGB for clarity. VGG16 was fed with these input images, each of size 224 x 224, of colours R, G, and B, respectively.

Face recognition: For face identification, an input face image was fed to the trained CNN to obtain its feature representation. This feature was compared with template features stored in a database to determine the person's identity. These features were extracted using HOG for all training samples. Once training was completed, test features were passed to the SVM too to predict the faces.

Classification Stage: This is where the model is defined, i.e., VGGNet is used as the foundational architecture for initialising our Convolutional Neural Network (CNN), and training takes place by using CNN to perform model classification. The feature vectors generated by the HOG descriptor were used to train the Support Vector Machines (SVM), and the results are validated against a given test input from both datasets.

Prediction stage: the trained CNN model was used to identify faces from the testing image data set.

Figures 2 and 3 show the flowcharts for both the MTCNN-VGG-16 and HOG-SVM models, respectively.

Figure 2
Workflow Diagram for CNN-VGG Base Model

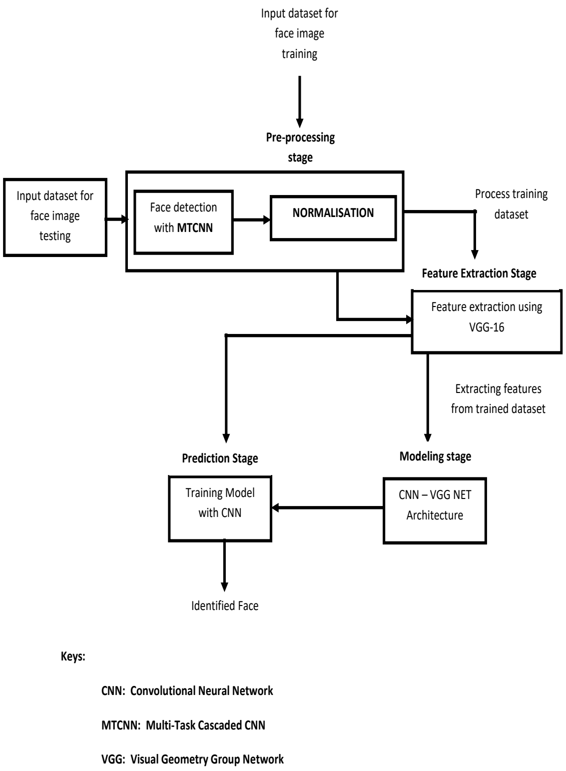
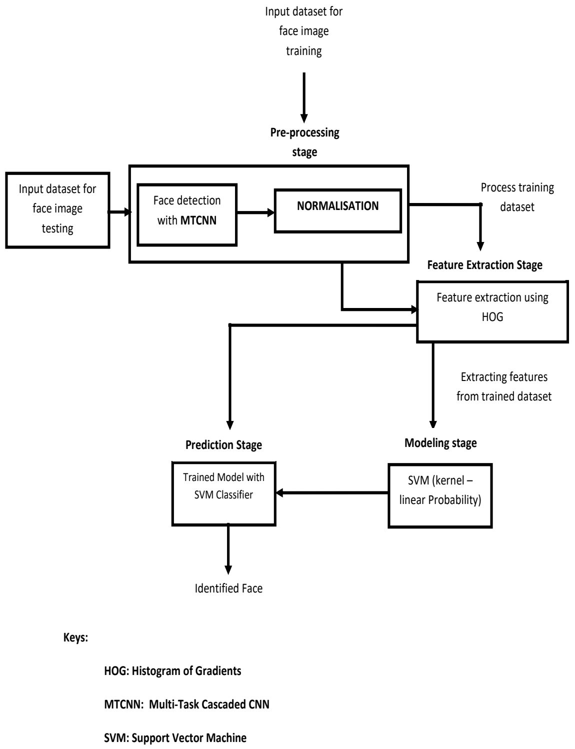


Figure 3
Flow Diagram for HOG with SVM Base Model



2.2 Dataset Structure

All face images are cropped and resized before being placed in separate folders for each class. A knowledge-based dataset is created by properly labelling the captured images with distinct classes.

2. 3 Implementation

A supervised learning approach was adopted for this study. HOG, the algorithm chosen for the feature extractor, was fed with the images of all faces in the dataset (occluded and non-occluded). 244 x 244 was the window size used since this size of face images corresponds to the aspect ratio of human faces in use generally. Blocks of pixels with 8 x 8 dimensions were used to calculate the descriptors. In turn, the descriptor values for each pixel over an 8 x 8 block are quantised into 9 bins, where each bin represents a directional angle of gradient and value in that bin—this value represents the combined sum of the magnitudes of all pixels with the same angle. The dimensions finally were 8 x 9 bins, which is equal to 72 feature length. This value was later changed from 16 x 16 and 32 x 32. Finally, histograms were used to train our Support Machine Classifier to detect and recognise faces in an input image.

The primary objective of all supervised learning frameworks is to determine the right label for fresh data coming in (see equation 1). In this equation, Y represents the predicted output, and x represents the input value. This function links input features to the predicted output of the trained model.

$$Y = (x) \qquad (1)$$

2.4 Support Vector Machine

For better accuracy in this study, SVM was also used to train the model. It was fed with 450 images divided into 15 classes, each with 30 images. The SVM performs the feature extraction, classifies extracted features, and generates a hyperplane that divides the discriminant features. In other words, suppose (Xi, Yi) is an extracted feature sample set with classification label Y = (+−11); then the hyperplane equation is as shown in equation (2):

$$(w.\, x) + b = 0 \qquad (2)$$

where w is a weight vector, x is an input vector, and b is a bias vector.

The optimal classification function of the SVM algorithm is:

$$g(x) = sign\left(\sum_{i}^{n} Y_i a_i^o K(X_i . X) - b_o\right) \qquad (3)$$

where, X_i is the support vector, a_i^o is the corresponding language coefficient, and b_0 is a threshold value.

From this equation, each discriminant feature is defined to be linearly separable case hence, by using non-linear transformation and continuously mapping high dimension space with the input vectors, the desired optimal hyper-plane will be realised.

In this case, the SVM is a polynomial classifier of degree d has the Radial function

$$K(X_i X) = exp\{-|X - X_i|^2/a^2\} \qquad (4)$$

And being a radial biased function, SVM serves as a Gaussian RBF (Radial Basis Function) classifier with Sigmoid function:

$$K(X_i X) = tanh\,(v(X_i . X) + C) \qquad (5)$$

Since SVM also act as a multi-layer perceptron in this case, then:

$Xi = (x1, x2, x3 \ldots xm)$ are the support vectors.

The K (Xi, X) indicates the kernel function.

And result of the SVM therefore is given as :

$$g(x) = sign\left(\sum_{i}^{n} Y_i a_i^o K(X_i . X) - b\right) \qquad (6)$$

All the features were extracted using HOG for all training samples and are then passed to SVM. The SVM classifier used a linear kernel to map the features. During training different regularisation values (C = {1e-10, 1e-5,1e-1,1,10,100,10000 and 1e6}) were tried. Once training was finished and the best C value had been determined for the classifier, the test features were then passed to the SVM to do prediction of the faces.

2.4 Performance Evaluation

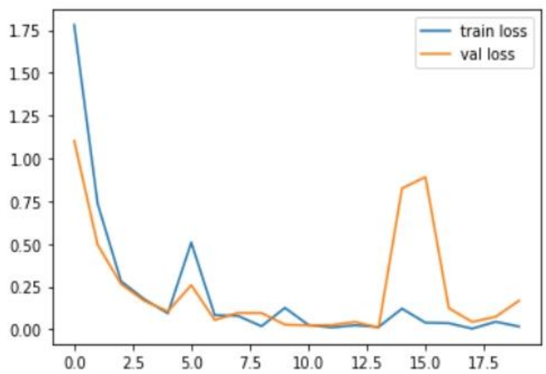
The following accuracy evaluation metrics chosen were based on the four possible outcomes: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

3.0 Results

In this study, a Convolutional Neural Network (CNN) based on the VGG-16 model was initially built to train and identify people with partial occlusion and without occlusion. Different epoch values were evaluated to achieve the best model accuracy. Each cycle a model takes through the entire training dataset is referred to as an epoch. In the experiments, 1 to 20 epochs were assessed on the training and validation datasets to obtain the best accuracy results of predicting the test data. Some of the results obtained at different epochs are shown in Table 1, and the plot of training and validation loss is illustrated in Table 2. Furthermore, the accuracy attained at different epochs is depicted in Figure 4 for both training and validation sets.

Figure 4

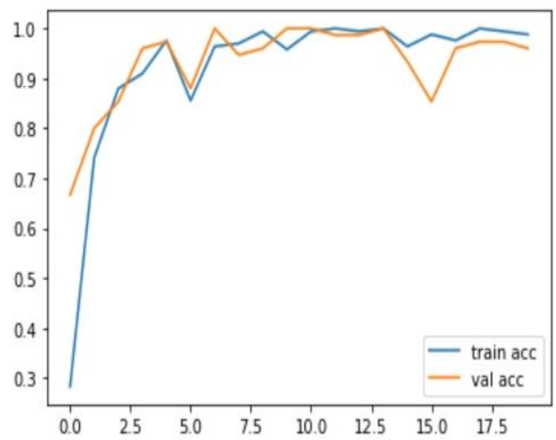
The Loss of the Model during Training and Validation



4.0 Discussion

Figures 4 and 5 represent the loss of the models on the training and validation sets, which mainly shows if the model is underfitting or overfitting. Since the decrease of the loss function for the training set is associated with a corresponding decrease in the validation data set too, this shows that this model needs no underfitting or overfitting, as it is a good fit.

Figure 5
Corresponding Changes in the Accuracy of Training Dataset and Validation Dataset



The accuracy of the training set is to determine if the model has been fully trained. From the results, it can be noticed that the training accuracy is almost 100%, which shows that the model has been fully trained. Starting from 30%, the accuracy on the validation set is also basically stable and improves steadily till it reaches >90%. From the results, it can be observed that the model achieved 95.6% accuracy after 20 epochs, which is the best result in this study using CNN.

4.1 Support Vector Machine

In the case of Histogram of Orientated Gradients (HOG) with SVM used mainly for face recognition with partial occlusions, several experiments were implemented to show the effectiveness of the SVM algorithm on the two datasets (partial occlusion and without occlusion). HOG was applied with different experimental parameters, and SVM was trained using different values of C. The results are as shown in Table 2 for different extract HOG parameters, grid, and regularisation parameters.

From the results in table 2, it can be observed that the performance of HOG and SVM increased steadily. Using a parameter of H=4 for HOG and (C = {1e-10 to 1e6}) for SVM, the best result is 62.2%. In contrast, using a parameter of H=8 HOG and (C = {1e-10 to 1e6}) for SVM, the best result is 72.72%. However, using a parameter of H=16 for HOG and (C = {1e-10 to 1e6}) for SVM, the best result is 79.54%. Further, with a parameter of H=32 for and (C = {1e-10 to 1e6}) for SVM, the best result is 84.9%.

This study proposed and compared two approaches for efficiently recognising individuals with partially occluded faces using Convolutional Neural Networks (CNN) based on VGG-16 architecture on the one hand and Histogram of Gradients (HOG) with Support Vector Machine (SVM) learning models on the other hand, running on the same set of data. Secondary data from Disguised Faces in the Wild and non-Caucasian nor Asian primary data from Africa were obtained and used to avoid bias. Using a fine-tuning strategy, optimisation techniques were specified in the training of the CNN model in order to improve its performance, taking advantage of pre-trained CNN. This CNN model realised an appreciable recognition accuracy of 95.6%.

Running on the same data sets, the HOG with SVM was also used mainly for face recognition in this study for both occluded and non-occluded images for comparison purposes. Along with SVM, HOG was applied to classify the results. This approach, which was tested on multiple images, produced a recognition accuracy of 84.9%. In addition, results show that the SVM algorithm can be regarded as a strong classifier for efficient dataset training, as it classified the tested data accurately, with a minimal error rate. It was also discovered that a larger training dataset will correspondingly improve the model's performance.

5.0 Conclusion

The development of a partially occluded face detection and recognition system using Convolutional Neural Networks (CNN) and Histogram of Orientated Gradients (HOG) with Support Vector Machines (SVM) holds immense potential in crime prevention. By addressing the challenges posed by partial occlusion, this system accurately detected and classified individuals with partially covered faces and can thus provide invaluable assistance to law enforcement agencies towards enhanced security measures.

Judging from several experiments conducted on the two models, the author concludes that both models (convolutional neural network (CNN) based on VGG-16 architecture and HOG with SVM) are useful and efficient in the detection and recognition of human faces that are either non-occluded or partially occluded. However, convolutional neural networks (CNN) demonstrated superiority in performance and accuracy over HOG with SVM in facial detection and recognition of partially occluded faces.

6.0 Recommendations

Stakeholders and security operatives in Nigeria and other nations facing crime control challenges are advised to invest in and adopt the outcome of this research in order to abate the soaring crime tide made possible because criminals can effectively hide their identity through face obliteration.

Future researchers can improve the performance of the model by training with larger datasets that include a diverse range of partially occluded faces for better accuracy. In addition, researchers are encouraged to explore other machine learning algorithms other than HOG with SVM and other variants of CNN deep learning algorithms on larger data sets. Besides, it is strongly recommended that datasets obtained from Africa be used in addition to others for training since criminal activities are not exclusive to Asian or non-African countries, and features peculiar to African faces cannot be substituted with non-African faces. This will remove elements of bias as well as make the model globally applicable.

7.0 Declaration of Conflicting Interest

There is no conflict of interest to declare for this study by the author.

8.0 References

Abbas, Q. (2021). FRS-OCC: Face Recognition System for Surveillance Based on Occlusion Invariant Technique. IJCSNS International Journalof Computer Science and Network Security, 21(8), pp. 288. Available: https://doi.org/10.22937/IJCSNS.2021.21.8.38

Adegoke, N. (2014). Kidnapping, Security Challenges and Socio-Economic Implications to the Niger Delta Region of Nigeria. Unilorin E-Journal (Online), 16(2), 205-216. Available at: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=CwWYrxMAAAAJ&citation_for_view=CwWYrxMAAAAJ:IjCSPb-OGe4C. [Accessed: 5, April, 2025].

Adegoke, N. (2020). Intelligence gathering and challenges of insecurity in Nigeria. African Journal of Criminology and Justice Studies 13(1), 4

Aladeselu, M. (2024). Boko Haram has caught the Nigerian Armed forces off-guard twice in less than two weeks in November 2024. – Zikoko news agency available at:https://www.zikoko.com › citizen › boko-haram-has-ca... [Accessed 6, March, 2025)

Ayegbusi, T. (2024). Armed Banditry and Kidnapping in Nigeria Advances in African Economic, Social and Political Development, in: J. Shola Omotola & Samuel Oyewole (ed.), The Political Economy of Kidnapping and Insecurity in Nigeria, chapter 10, pp. 105-134, Springer.

Bachhety, S., Singhal, R., Rawat, K., Joshi, K., & Jain, R. (2018). Crime Detection Using Text Recognition and Face Recognition. ttp://www.acadpubl.eu/hub/

Bansal, A., Nanduri, A., Castillo, C. ., Ranjan, R., & Chellappa, R. (2018). An annotated face dataset for training deep networks. In 2017 IEEE international joint conference on biometrics (IJCB), 464-473). IEEE.

Chethana, H. ., Trisiladevi, C. ., and Shashank, M. . (2021). Detection of Partially Occluded Faces Using Convolutional Neural Networks. In 3rd International Conference onIntegrated Intelligent Computing Communication & Security (ICIIC 2021), 69-76. Atlantis Press.

Kleffmann, J., Ramachandran, S., Cohen, N., O'Neil, S., Bukar, M., Batault, F., & Broeckhoven, K.,. Banditry Violence in Nigeria's North West: Insights from Affected Communities, Findings Report 36, UNIDIR, Geneva, 2024, https://doi.org/10.37559/MEAC/24/05

Kortylewski, A., Liu, Q., Wang, H., Zhang, Z., & Yuille, A. (2020). Combining Compositional Models and Deep Networks For Robust Object Classification under Occlusion.

Kostomarov, K. . (2019). Possibilities of using neural networks in the investigation of crimes. *Journal of Siberian Federal University. Humanities & Social Sciences,* 12(11), 2023–2030. https://doi.org/10.17516/1997-1370-050

Nakib, M., Hasan, M. S., Khan, R. T., & Uddin, J. (2018). Crime Scene Prediction by Detecting Threatening Objects Using Convolutional Neural Network.

Nigeria Violence: Governor's Office Revises Previous Toll of 17 In Benue State | WION, (2025) Available at: https://www.youtube.com/watch?v=nG4VAhwFAss. [Accessed on 6 May, 2025]

Prabakaran, S., & Mitra, S. (2018). Survey of Analysis of Crime Detection Techniques Using Data Mining and Machine Learning. Journal of Physics: Conference Series, 1000(1). https://doi.org/10.1088/1742-6596/1000/1/012046

Sreelakshmi P, & Sumithra M. (2019). Facial Expression Recognition robust to partial Occlusion using MobileNet. www.ijert.org

Surajit S. (2017). Image Analysis and Processing - ICIAP 2017 (S. Battiato, G. Gallo, R. Schettini, and F. Stanco, Eds.; Vol. 10485). Springer International Publishing.https://doi.org/10.1007/978-3-319-68548-9

Valueva, M., Nagornov, N. Lyakhov, P., Valuev, G., & Chervyakov, N. (2020). Application of the residue number system to reduce hardware costs of the convolutional \ neural network implemen tation. Mathematics and computers in simulation, (MATCOM) Elsevier, 177(C), 232-243.

Vivian, N. ., & Ise, O. . (2020). Face recognition service model for student identity verification using deep neural network and support vector machine (SVM). Int J Sci Res Comput Sci Eng Inf Technol, 6(4), 11-20.

Ge,Y., Huicheng, Z., & Jiayu D. (2022). MSML: Enhancing Occlusion Robustness by multi-scale segmentation-based mask learning for face recognition . The Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-22) 3197-3205 https://cdn.aaai.org/ojs/20228/20228- 3-24241-1-2-20220628.pdf

Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., & Liu, W. (2018). Cosface: Large margin cosine loss for deep face recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition . 5265-5274).

Wang, M., Deng, W., Hu, J., Tao, X., & Huang, Y. (2019). Racial faces in the wild: Reducing racial bias by information maximization adaptation network. In Proceedings of the IEEE international conference on computer vision,. 692–702

Wang, X., Zhang, S., Wang, S., Fu, T., Shi, H., and Mei, T. (2020). Mis-classified vector guided softmax loss for face recognition. In Proceedings of the AAAI Conference on Artificial Intelligence, 34(17), 12241-12248.

Yang, L., Ma, J., Lian, J., Zhang, Y., & Liu, H. (2018). Deep representation for partially occluded face verification. Eurasip Journal on Image and Video Processing,2018(1). https://doi.org/10.1186/ s13640-018-0379-2 [Accessed: 17, Nov. 2024]

Zhao, X., & Wei, C. (2017). A real-time face recognition system based on the improved LBPH algorithm. In 2017 IEEE 2nd international conference on signal and image processing (ICSIP) (pp. 72-76). IEEE.

Zhao, B., Lu, H., Chen, S., Liu, J., & Wu, D.

(2017). Convolutional neural networks for time series classification. Journal of Systems Engineering and Electronics, 28(1),. 162-169. Thirteenth Report on Violence in Nigeria. (2023) [Accessed on: 6, May, 2025]