American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-11, Issue-11, pp-16-25
www.ajer.org
Research Paper Open Access

Image Pattern Recognition Using Deep Learning Model

[1]OJO O.S.A [2]MBACHU C.B. [3]ILOH J.P.I.
[1]Electrical/Electronic Engineering Department., Federal Polytechnic Oko, Anambra State
[2,3]Electrical/Electronic Engineering Department., Chukwuemeka Odumegwu Ojukwu University, Uli, Anambra State

Abstract: *The rapid emergence and evolution of deep learning techniques and the increasing availability of data have allowed AI researchers to solve many problems that had been previously unsolvable. The advent of comprehensive frameworks for developing AI has also opened the door to the development of new tools and techniques in applied fields.Through information systems research, we can bridge the gap between the theoretical and practical aspects of image recognition.As a result of the increasing importance of pattern recognition in research, various techniques have been developed to improve its performance. One of these is Deep Learning.*

This paper presents a novel deep learning algorithm using convolutional neural network to provide convolutional face recognition algorithm which is reliable to process all forms of image data types irrespective of the noise level attributed to it and achieve quality recognition result with high precision.

Keywords: Image recognition, deep learning, artificial intelligence, object detection, neural networks, machine learning

---

Date of Submission: 02-11-2022 Date of acceptance: 13-11-2022

I. INTRODUCTION

Artificial intelligence (AI) researchers have been able to develop new techniques and solve previously unsolvable problems due to the increasing number of data sources and the technological advancements that have occurred over the past few years. Artificial intelligence approach remains the most reliable means of solving pattern recognition problem, due to its ability to learn from training dataset and then makes accurate decisions. For instance, deep learning has been able to reveal intricate structures in complex image and voice recognition systems. Due to the immense amount of data collected by researchers, deep learning applications can be widely used in various fields, such as healthcare, business intelligence, marketing, science etc.

Data mining has been widely used in the development of advanced models, which significantly improves the efficiency of their implementations. Pattern recognition is a process that involves extracting hidden patterns from large datasets. It is a study of how machines can learn to distinguish between various types of objects and their backgrounds. After identifying the appropriate objects, the system then makes informed decisions regarding their output.

Recently, deep learning has been widely used in the development of efficient and effective pattern recognition systems. It is mainly used in machine learning to model complex data structures using different types of transformations. Deep learning falls within the domain of neural networks, in which neural networks have more than three layers, i.e. more than one hidden layer. These neural networks used in Deep learning are called Deep Neural Networks (DNNs). DL algorithms are similar to how nervous systems are structured, where each neuron is connected to each other to pass on information. DL models work in layers and a typical model at least have three layers and each layer accepts the information from previous and pass it on to the next one (Levine et al., 2016). It is a technique that allows you to learn in a deep way without a specific code for it.

Deep Learning fundamentally performs much more advanced functions, allowing the analysis of a wide range of factors at the same time. For example, Deep Learning is used to contextualize the information received by the sensors used in autonomous cars: the distance of objects, the speed at which they move, predictions based on the movement they are making, etc. They use this information to decide how and when to change lanes, among other things (Teoh et al., 2021).There are two main types of training methods in deep learning: unsupervised and supervised.

In supervised learning, the model learns from a labeled training dataset. On the other hand, in unsupervised learning, the model looks for correlated patterns in the data set to find the hidden structure.
In this paper, we present a novel deep learning algorithm that can be used to perform face recognition based on various image data types. It is capable of achieving high accuracy and reliability even with the varying noise levels generated by the algorithm.

II. LITERATURE SURVEY

Praneeth et al. (2020) presented research on face recognition system based on deep learning technique. In the study, the features of face were extracted using the binary pattern histogram algorithm and then trained with convolutional neural network. The performance can be improved using image processing technique.

Micheal and Aman (2020) presented a research on face recognition using deep learning technique. In the study, a systematic review was performed, and then convolutional neural network was adopted for the training faces with an accuracy of 97% when deployed and tested on tensor-flow software. The performance was however limited to the training dataset used.

Marios (2018) researched on a simultaneous low resolution an off pose angle face matching algorithm as an investigative lead generative tool for law enforcement. The study presented a security system which has the capacity to detect faces of suspected individuals who disguised themselves to commit crime. The system when tested showed good accuracy on multiple datasets, but can be improved using deep learning algorithm.

Nishtha (2019) studied face recognition using deep convolutional neural network. In the research the architecture was developed and used to train ORL dataset which has 40 classes of different peoples. The result when tested showed that the accuracy is 95%; however, the performance when deployed in real time application may not be accurate as the data used to achieve this 95% accuracy is already processed. Hence there is need to incorporate filter to the system to sustain the same level of accuracy when deployed in real life application.

Zeelan et al. (2020) researched on face mask detector with Pytorch and Deep learning techniques. The study used OpenCV, artificial neural network (ResNet) and computer vision to develop a system which could detect if a person is wearing face mask. The aim was to prevent the spread of coronavirus. The percentage accuracy when tested was 97%; however, the test was not done with a standard dataset, and hence could not be justified.

Prajyot (2020) researched on Deep learning approach for facial expression recognition. In the study, deep learning algorithm was developed and deployed for the recognition of face expression using convolutional neural network technique. The performance was evaluated using confusion matrix on Jaffe dataset and the performance was efficient. However, despite the success, the performance could be improved using image processing.

Aishwarya and Anirban (2020) researched on multi target facial recognition system using deep learning technique. In the study, it was revealed that the main problem of face detection system was heavy computation time and limited reusability. The study utilized computer vision, image processing, deep learning, Google faceNet and sample dataset containing 1930 images. The performance when tested showed high accurate 98.7% and response time of 11.50s. However despite the success, it is clear that the delay time will increase on a more robust dataset and hence the need for improvement.

Ayusha et al (2019) presented a deep learning based cardless ATM system using fingerprint and face recognition system. In the study, convolutional neural network was designed and trained with face and fingerprint dataset for securing automated teller machine. The hybrid biometric technique idea was welcomed as it improved the security of the system, but the performance could be made reliable using image processing.

III. METHODOLOGY

The methods used for the development of the proposed system are data collection, image acquisition, computer vision, face detection, data pre-processing, data processing, training and face recognition.

3.1 Data Collection

For the AI to be properly trained in the proposed system, the primary source of data collection is the Nigerian Police Force (NPF), Central Police Station (CPS), Awka, Anambra State. The data so collected formed the datasets used in this research including secondary sources such as oral interviews, emails, investigation etc.

3.2 Data Processing

The dataset provided by the NPF has a sample size of 1850 different suspected identities apprehended for various reasons and awaiting trials. The data were already sized and formatted according to their specification of the NPF as presented in table 1 and stored for official use.To perform comparative analysis for evaluation of proposed work.

Table 1: Data properties

Image parameters	Values
Dimension	190 x 260
Height	260
Weight	190
Horizontal Resolution	96dpi
Vertical resolution	118dpih
Bit depth	24
Size	5.6kb

3.3 Image Acquisition

To enhance the accuracy of the recognition process and due to requirement of high-quality image, a high-definition camera was used to capture the image. However, due to the fact that HD camera may not be available at the site of crime for capturing of quality images, computer vision is required as discussed below.

3.4 Computer Vision

In computer vision, the camera is incorporated with intelligence which enables it to derive useful from a digital image before capturing. There are various computer vision algorithms in (Girshick et al., 2014) which can achieve this objective today. However, the adoption of any algorithm is dependent on the type of data under investigation, which in this case is a human face.

In this study, the Viola Jones algorithm was used because it enables the camera to detect the face and also tracks it in real time to collect quality data. According to Vilmal (2015), the algorithm has many advantages, enables via its sophisticatedandinvariant detector which locates the scales of human facial features. This algorithm was reconfigured to be compatible with the specifications presented in table 1. The aim is to ensure that any image captured in time series are of the same size and properties. The data flow model of the reconfigured Viola Jones algorithm is presented in fig.1 as flowchart.

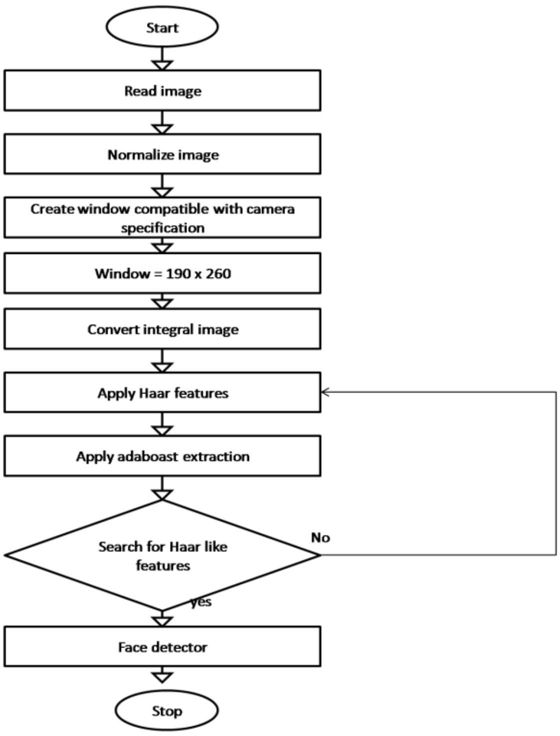

Fig.1: The Configured Viola Jones algorithm

The flowchart in fig.1 presented as algorithm shows how it was configured to align with the properties of the recommended camera. The algorithm came into play when the camera reads image data using Haar features and Adaboast algorithms and then recognizes facial fudicial points (features) to detect a face.

3.5 Face detection

This is the process of identifying and capturing of face using computer vision technology. The camera reads a computer vision-based scales image according to Haar features in fig. 1 and then acquire the data. In this case the camera, despite its ability to view other objects within its focus, identifies only face-like images and captures it.

3.6 Data pre-processing

The data pre-processing commonly involves removing low-frequency background noise, normalizing the intensity of the individual particles images, removing or enhancing data images prior to computational processing; and this was done using Histogram Equalization Technique (HET) as specified in (Christophe et al., 2007). HET is a contrast adjustment technique developed as a low band pass filter to eliminate unwanted noise attributed with low frequency and was adopted to pre-process the real image.

3.7 Data processing

The impulse noise and stripped noise which cause severe degradation on the image quality according to Mohd et al. (2012), has to be processed to guarantee quality image result. The aim is to ensure that the most interesting features of the face were made available in good quality for optimized training and recognition result. To achieve this, wavelet filtering and normalization techniques were used.

3.8 Training Strategy

Training artificial intelligence is a technique used to learn a machine learning or deep learning algorithm of particular data patterns, and the reference model learning used for time series classification decisions. In this study, the technique was used to train the dataset images collected so as to learn the pattern of each image as a reference convolutional face model and used for time series face recognition. This training process was done using convolutional neural network. The convolutional neural network is a deep learning algorithm which specializes in solving image-based pattern recognition problems and it uses the images to develop a facial recognition system.

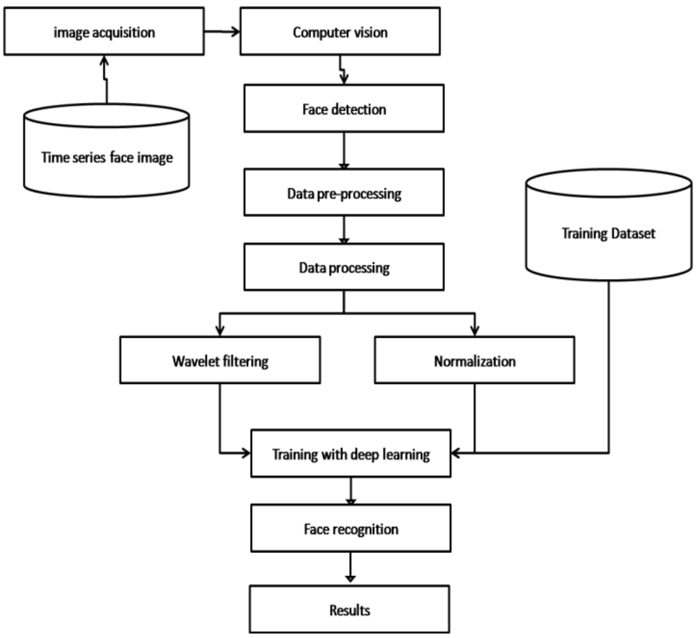

Fig. 2: Process Diagram

The process diagram was used to present the logical interaction of the various methods discussed to achieve the proposed system as shown in fig. 2.

The process diagram presented the face detection and recognition system workflow. The diagram shows how the training dataset containing images of suspected persons was trained to generate a reference model, and then use it to recognize time series face images. This was done using image acquisition process which intelligently collects the image of the suspect with the help of computer vision technology. The computer vision was used to make the camera only search for face-like images, based on Haar features and Adaboast algorithm to detect face and capture it. The data captures which is always characterized with noises; such as background noise, strip noise and impulse noise were processed using histogram equalization technique to remove the background noise; wavelet filter to remove the impulse noise and then normalization process to mitigate the strip noise. The processed data was then fed to the deep learning algorithm for training using convolutional neural network which train the input data and compare with the reference model to recognize face.

3.9 The Proposed Deep Learning Model

The deep learning model was developed using Convolutional Neural Network (CNN). This is deep learning algorithm which has proven in other areas of application to be the best in solving image-based pattern recognition problem. This was modeled in this research to train the wavelet image processed and to recognize face accurately. The CNN is made up of input layer responsible for the dimensioning of the image into a compatible size and then fed to the convolutional layer for scanning and extraction of the best feature parts, however these high quality of image pixels scanned, pooled and arranged in the last convolutional layer contains noise reflection originating from the high feature pixels which requires normalization before feeding to the fully connected layer. This was done using rendering technique to normalize the pixels behaviors and then fed to the fully connected layer for training using artificial neural network and then predict the output with softmax function. The block diagram of the CNN is presented in fig. 3.

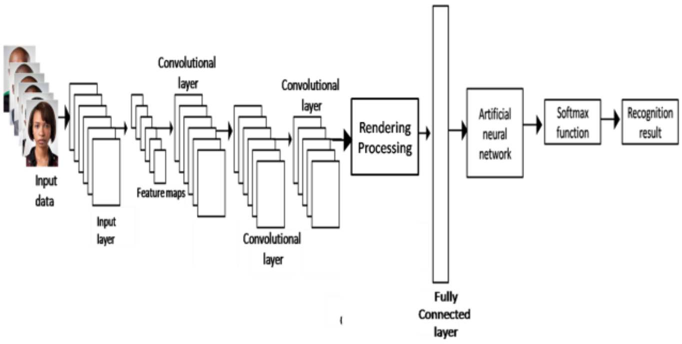

Fig. 3: The CNN Block Diagram

The input layer of the CNN collects data from the wavelet transform in the form of dimension (hxWxC). The dimension represents height, weight and color channel respectively. The input dimension fed to the CNN is (190x260x3).

Furthermore, the various pixels of the dimension image is extracted using interesting key features of the image in a pixel format. The filter size was specified as:

F = (fw x fh x d) (1)

Where fw is the filter weight, fh is the filter height and d is color dimension. A matrix is thereafter formed from the scanned pixel information. The model of the feature maps scanned by the filters is presented in equation 2.

$$F_o = \left[\frac{F_i + 2_p - k}{s}\right] + 1 \qquad (2)$$

Where F_o is the output features, F_i is the input features, p is the convolutional padding, s is the strides size, k is the convolutional kernels size. During these feature extraction process, rectified linear unit (ReLU) was used to introduce nonlinearity in the features to transform all negative values into null. The next stage was to pool the features into matrix presentation of a convolutional layer. The image pixels collected during scanning were pooled using maximum pooling technique to select the highest pixel value of array of extracted pixel matrix to form a convolution layer as shown in fig. 4.

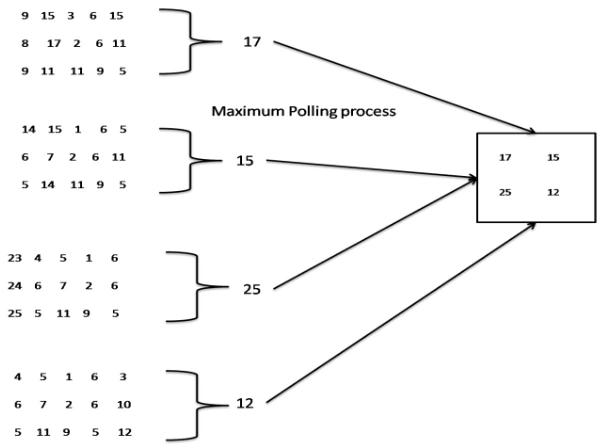

Fig. 4: The Maximum Pooling Process

The fig. 4 presents some of the output 5 by 3 matrixes of pixels scanned by the 5 by 3 filter determined by equation 1. From the output as shown in the matrixes above, the highest values were only selected and used to form the next convolutional layer. The reason is that the pixel with the highest value contains the most important part of that image section while the total output was presented as a convolutional layer as shown in fig. 5

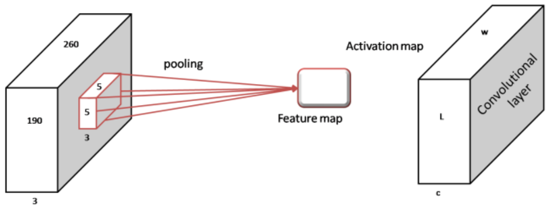

Fig. 5: The Convolutional Process

The fig. 5 presents the data flow diagram of the convolutional layer setup. This was achieved using the 5 by 3 filter to scan the 190 by 260 face image. The scanned output was presented as feature map and pooled into a matrix of pixels to form a convolutional layer. The total pixel used to form the convolutional layer is presented as equation 4.

$C_O=((w*h) + 1)*nf)$ (3)

Where the total pixels in the convolutional layer is given as (C_o), w is filter weight (w), filter height (h), number of filters (nf) and filter bias term (1). The model in equation 3 presents the total pixels values in the first convolutional layer only; however, to generate pixels for other convolutional layer, the number of previous pixels was added (np) and presented as:

$C_O=((w*h * np) + 1)*nf)$ (4)

This model was used to sum up all the feature maps extracted per convolutional process in an array of matrix. The activation size is determined using the relationship between the number of image pixels and the depth as shown in equation 5.

$A_s=(w*h*d)$ (5)

In this model of equation 5, the parameters of w, h and d are defined based on each layer specifications.

3.10 Rendering Model

The rendering model is presented using the relationship between the radiance reflected from each image pixel and reflection distribution factor from all angles of feature points as presented in equation 6.

$dL_r(x,y) = f_r(x,w,y)L_f(x,w)\cos\theta dw$ (6)

Where x is the total reflectance from each image pixel, w is the convolutional angle, and θ is the angle of the feature point, dw is the differential angle of the pixel, L_f is the surface, y is the direction of radiance, L_r is the surface radiance.

The rendering model in equation 6 was applied after the last convolutional layer of the CNN considering the emitted radiance from the image pixel and presented as an integral rendered convolutional image as equation 7.

$L_s(x,y) = L_e(x,y) + \int f_r(x,w,y)L_f(x,w)\, cos\,\theta dw$ (7)

Where radiance on the convolutional image at the last convolutional layer is Ls, radiance emitted from the image feature is presented as Le,

The equation 7 can be remodeled as equation 3.13.

$$L_s(x,y) = L_e(x,y) + \int g(x,k)f_r(x,w,y)L_f(x,w)\frac{cos\,\theta\, cos\,\theta^i\, dA}{||k-x||^2}$$

(8)

Where $||k-x||^2$ is the distance from k to x; $\cos\theta^i = (n.w)$; $L_s(x,y) = L_e(x,y)$ and

$g(x,k)= \begin{bmatrix} 1 & if x is visible to k \\ 0 & otherwise \end{bmatrix}$

3.11 Loss Function Model

This section of the CNN flattens the rendered convolutional image in equation 7 and then fed to a neural network for training to learn the image features for time series face recognition. The training was enabled using back propagation algorithm in fig. 6, and was monitored for accuracy to ensure that there is no overshoot performance using the loss function model in equation 9.

$L = \sum_{i=l}^{K}(P_i - D_i)^2$ (9)

Where L is the loss function, k is the number of observations, P is prediction and D is the training target. The model presents the loss between P and D, where K is the number of class in the face dataset and i is the output loss which is scalar. The Deep Learning settings are shown in Table 2.

Table 3.2 Deep Learning Settings

Maximum number of epoch to train	15
Epoch between display	1
Maximum time to train in sec	Infinity
Maximum validation failure	5
Scale factor for length	190
Scale factor for weight	260
Initial step size	0.01
Minimum performance gradient	1e-6
Cost horizon	7
Control horizon	2
Number of bias function	1
Number of Channel	3

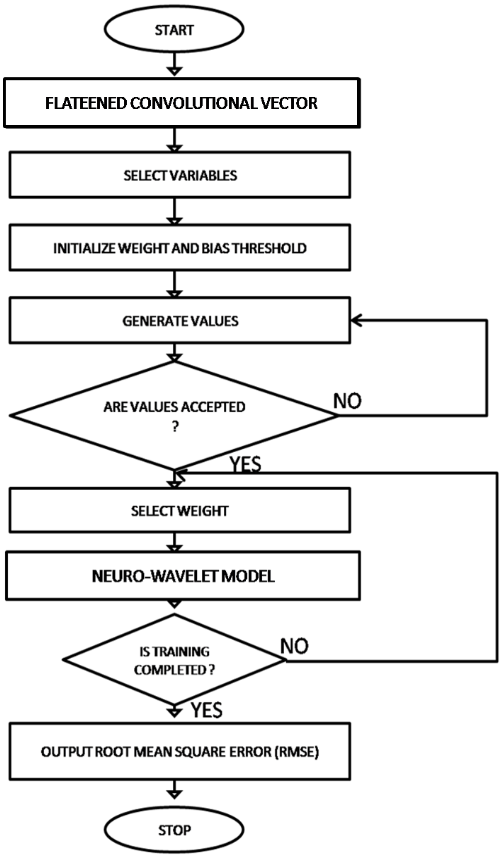

Fig. 6: The Back Propagation Algorithm

3.12 Output Layer for Recognition of Result

This is the final layer of the network which produces the desired output of the training process. This layer is designed using a softmax activation function which transforms the learned feature vectors into probability distributions consisting of various probabilities proportional to the various exponential of the face input data. The model of the softmax function is presented in equation 10.

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \qquad (10)$$

Where σ is the softmax function; e^{z_i} is the standard exponential function for input vectors from the ANN output, k is the number of classes of the multiset classifiers; e^{z_j} is the standard exponential function for output vector, $\vec{z}$ is the input vector.

An improved CNN modeling diagram is shown in figure 7.

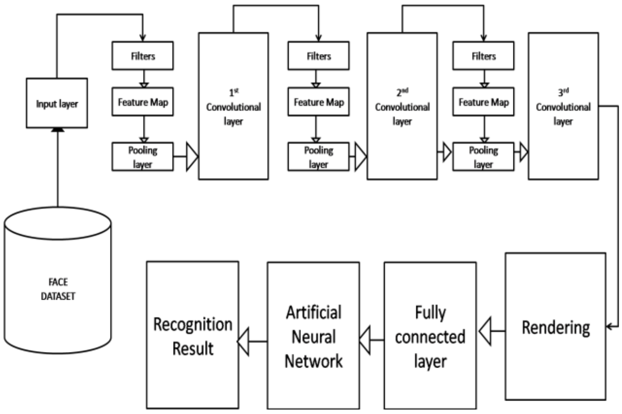

Fig. 7: An Improved CNN modeling diagram

The fig. 7 presents the model of the novel convolutional neural network developed with the potential for précised facial recognition result. From the diagram, the face dataset is fed to the CNN via the input layer for dimensioning and then scanning using filter and pooling layer. The outputs are arranged in matrix array to form a convolutional layer and then the scanning process was repeated until the third convolutional layer. At this stage the rendering process was introduced for normalization before feeding to the fully connected layer which uses artificial neural network to train the data for time series facial recognition purposes. The system flow chart is presented in fig. 8

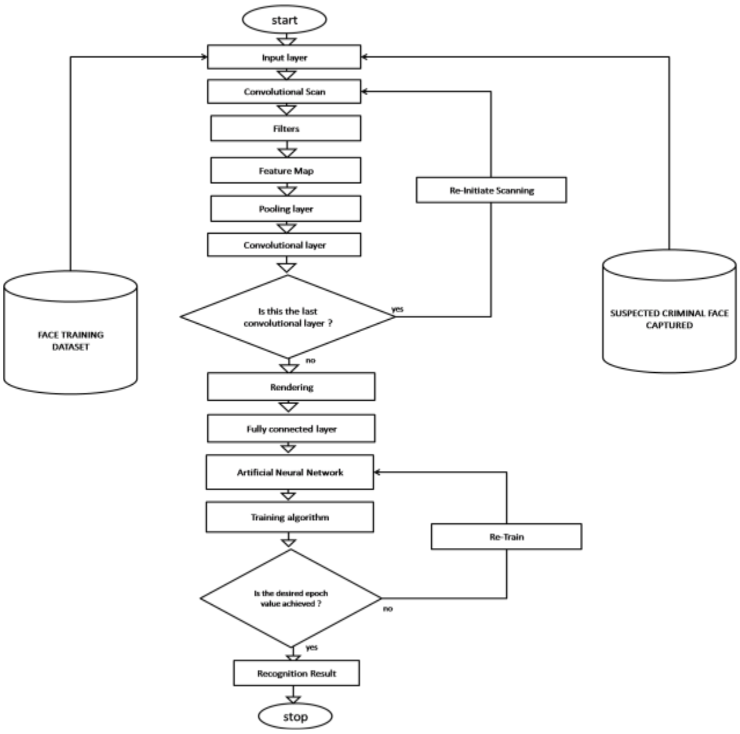

Fig. 8: System Flowchart

The system flowchart presents the data flow activity of the new system developed, showing how the training and test data was processed by the system to make accurate recognition process. From the flowchart, the convolutional neural network is first trained and used to recognize faces of suspected and captured person. The convolutional neural network model was developed using rendering approach which normalizes the convoluted series of high value pixels and then used artificial neural network to train the data for recognition result.

IV. RESULTS

The performance of the CNN algorithm was measured using the loss function model in equation 9. This model was used to evaluate the accuracy of the training process via iterative evaluation of epoch at each training steps. To perform this training process, the dataset collected the NPF was loaded into the CNN algorithm developed in fig. 6 and then trained using deep learning toolbox configured with the table 2 settings.

Before the training began, the feature vectors were automatically divided into three sets which are: training, testing and validation sets in the ratio of 70:15:15. The aim is to evaluate the CNN algorithm by simultaneously training, testing and cross validating the performance as shown in the fig. 9

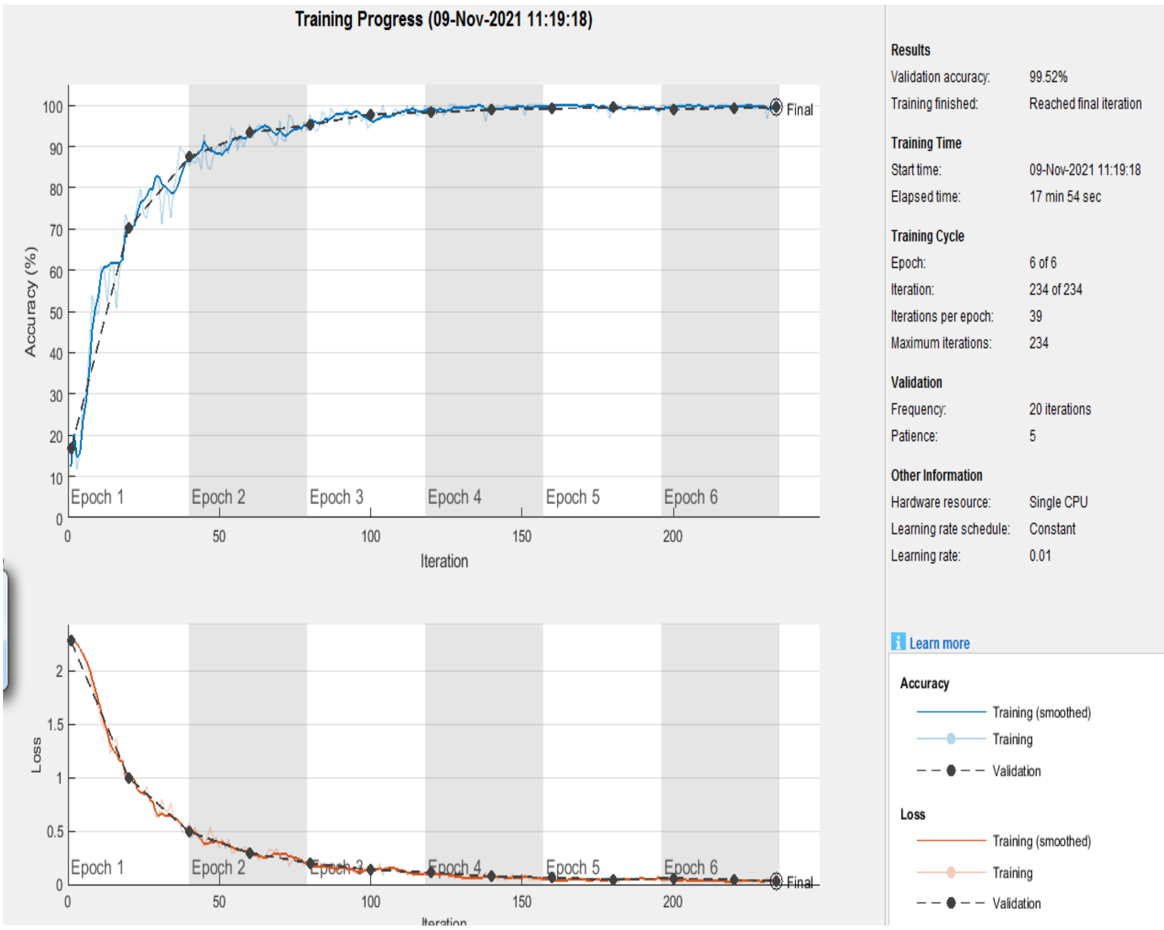

Fig. 9: Deep Learning Algorithm Result

The fig. 9 presented the performance of the CNN Algorithm developed in the model of fig. 6 and measured with the Loss function presented in equation 9. The result showed the relationship between the multi sets and their performance at six different epoch cycles. From the result, the overall training performance was measured which computes the mean accuracies between the loss function for the training, testing and validation sets and the result is 99.52%. The loss function is 0.48% which is very good as it indicated a very acceptable Mean Square Error (MSE) result.

The implication of the result showed that the CNN algorithm was very efficient in learning the data of the criminals and ready for time series facial recognition with guaranteed high accuracy. This result was validated using tenfold cross validation technique presented in equation 11.

$$CVA = \frac{1}{10}\sum_{1}^{10} Ai \quad (11)$$

Where CVA stands for Cross Validation Accuracy, A is the accuracy measure for each fold; the results are presented in table 3.

Table 3: Ten-fold Training Results

Training times	Accuracy (%)	Loss (%)
1	99.45	0.55
2	98.91	1.09
3	99.44	0.56
4	99.51	0.49
5	99.22	0.78
6	99.60	0.40
7	98.91	1.09
8	99.47	0.53
9	99.39	0.61

10	98.90	1.10
Average	99.28	0.72

The results in table 3 presented the Tenfold performance of the CNN algorithm using the model in equation 11. The average was computed and the result showed that the accuracy of the CNN algorithm is 99.28% which is very good. The loss function was also measured and the result showed 0.72% which is also very good, showing minimum error which is acceptable and negligible.

Having trained the algorithm, tested and validated its effectiveness to accurately recognize criminals, the algorithm was deployed as facial recognition software and tested.

V. CONCLUSION

The result of CNN algorithm as presented shows the performance of the algorithm when implemented as a facial recognition system, this was tested using query face, using face detection model in fig. 2 Using Viola and Jones algorithm fig.1 the developed model enhanced the accuracy of face detection. From the result obtained it was observed that the face was accurately detected.

Final result of the queried imageshows how the algorithm was used to search and detect the face of the suspect from the training dataset.

REFERENCES

[1]. Aishwarya Ramesh, AnirbanGhosh (2020)" Multi Target Facial Recognition System Using Deep Learning" IOSR Journal of Computer Engineering (IOSR-JCE)e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 22, Issue 4, Ser. I , PP 01-07www.iosrjournals.org

[2]. Ayusha Mohite, Shruti Gamare, Karan More, Nita Patil (2019)"Deep Learning based Card-Less Atm using Fingerprint and Face Recognition Techniques" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056; Volume: 06 Issue: 03 | Mar 2019 www.irjet.net p-ISSN: 2395-0072

[3]. Christophe Garcia, Jörn Ostermann, and Tim Cootes (2007)"Facial Image Processing" EURASIP Journal on Image and Video Processing" Hindawi Publishing Corporation Volume 2007, Article ID 70872, 2 pages; doi:10.1155/2007/70872

[4]. Girshick, J. Donahue, T. Darrell, and J. Malik (2014) "Rich feature hierarchies for accurate object detection and semantic segmentation". In CVPR.

[5]. Levine, S., Finn, C., Darrell, T. &Abbeel, P. (2016). "End-to-end training of deep visuomotor policies," J. Mach. Learning Res., vol. 17, no. 39, pp. 1–40.

[6]. MariosSavvides, (2018) " A Simultaneous Low Resolution and Off- Pose Angle Face Matching Algorithm as an Investigative Lead Generative Tool for Law Enforcement", U.S. Department of Justice.

[7]. Michael Farayola, Aman Dureja (2020) "A Proposed Framework: Face Recognition With Deep Learning" INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 9, ISSUE 07.

[8]. Mohd B., Sandeep S., Nelson K (2012) " Survey on noise removal in digital images" IOSRJCE; ISSN 2278-0661; vol 6; Issue 4; PP 45-51

[9]. NishthaJatana (2019)"Face Recognition using Deep Convolution Neural Network" IJEDR 2019| Volume 7, Issue 4 ISSN: 2321-9939

[10]. Prajyot H. Mohite, Pratibha P. Shingare (2020)" Deep Learning Approach For Facial Expression Recognition" Journal of critical reviews;ISSN- 2394-5125 Vol 7, Issue 14.

[11]. Praneeth T, Rajesh K, Naga Raju, Suneetha Manne (2020) "Face Recognition Based Attendance System Using Deep Learning Techniques" Journal Of Critical Reviews ISSN- 2394-5125 VOL 7, ISSUE 12.

[12]. Teoh K.H, RC Ismail, SZM Naziri, R Hussin, MNM Isa and MSSM Basir (2021)" Face Recognition and Identification using Deep Learning Approach" 5th International Conference on Electronic Design (ICED;)Journal of Physics: Conference Series 1755 012006IOP Publishing doi:10.1088/1742-6596/1755/1/012006

[13]. Vimal, Virender Kadyan (2015). Implementation and Performance Analysis of Face Recognition Using MATLAB. Dept. of Computer Science and Engineering, RNCET, Panipat, Haryana,. IJCST Vol. 6, Issue 2, ISSN : 0976-8491 (Online) | ISSN : 2229-4333.

[14]. Zeelan Basha1, B.N. Lakshmi Pravallika and E. Bharani Shankar (2020)" An Efficient Face Mask Detector with PyTorch and Deep Learning CMAK" EAI Endorsed Transactions on Pervasive Health and Technology 12; Volume 7; Issue 25

OJO O.S.A, et. al. "Image Pattern Recognition Using Deep Learning Model."*American Journal of Engineering Research (AJER)*, vol. 11(11), 2022, pp. 16-25.