Journal of Information Systems and Informatics
Vol. 6, No. 1, March 2024 e-ISSN: 2656-4882 p-ISSN: 2656-5935
DOI: 10.51519/journalisi.v6i1.635 Published By DRPM-UBD

Predictions of Criminal Tendency Through Facial Expression Using Convolutional Neural Network

James, Gabriel Gregory[1], Okafor, Peace Chiamaka[2], Chukwu, Emenike Gabriel[3], Michael, Nseobong Archibong[4], Ebong, Oscar Aloysius[5]

[1]Department of Computing, Faculty of Computing and Applied Sciences, Topfaith University, Mkpatack, Nigeria
[2]Department of State Services, Bayelsa State Command, Bayelsa State, Nigeria
[3]Department of Software Engineering, Federal University of Technology, Ikot Abasi, Nigeria
[4]Department of Mathematics and Computer Sciences, Ritman University, Ikot Ekpene, Nigeria
[5]Netisens ICT Hub, 133 Atiku Abubakar Way, Uyo, Akwa Ibom State, Nigeria
Email: [1]g.james@topfaith.edu.ng, [2]okaforpeace@gmail.com, [3]chukwugabriel0010@gmail.com, [4]nsemike@yahoo.com, [5]oskndott@gmail.com

Abstract

Criminal intention is a critical aspect of human interaction in the 21st-century digital age where insecurity is on the high side as a major global threat. Kidnapping, killings, molestation of all sorts, gender-based violence, terrorism, and banditry are the trends of criminality in our nation, as such, there is a need to effectively explore innovative means to identify and cope with this evil menace in our society. The facial positioning of humans can tell their evil intention even if they pretended to smile with the evil in their minds. In normal instances, it may be very difficult to predict the heart of man, but with the trending information technology like image processing, the state of a human face could be used as a means to read their tendencies. This paper proposes a deep learning model based on the FER2013 dataset through the implementation of a CNN model that predicts criminal tendencies with the help of facial expressions. With this goal in mind, we explore a new level of image processing to infer criminal tendency from facial images through a convolutional neural network (CNN) deep learning algorithm in other to discriminate between criminal and non-criminal facial images. It was observed that CNN was more consistent in learning to reach its best test accuracy of 90.6%, which contained 8 convolutional layers. To increase the accuracy of this model, several procedures were explored using Random Search from the Keras tuner library, testing out various numbers of convolutional layers and Adam optimizer. It was also noticed that applying the dissection and visualization of the convolutional layers in CNN reveals that the shape of the face, eyebrows, eyeball, pupils, nostrils, and lips are taken advantage of by CNN to classify the images.

Keywords: Criminal Tendency, Facial Expression, Convolutional Neural Network, Machine Learning.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

1. INTRODUCTION

Tendency is an inclination towards a particular characteristic or type of behavior. It is a proneness to a particular thought or action. There are four categories of tendencies: Upholders, Questioners, Obliges, and Rebels. These categories of the premise of expectations, specifically, how one responds to internal and external expectations.

There have been confirmed incidences of kidnapping and murder in our culture in the past and even currently. Today's society is rife with dreadful things like the ones listed above. By interacting with people and working to alter behavior, we have attempted to lower the frequency of this crime in our society. One of the biggest whys is murder. Why was she killed by him? Alternatively, why harm defenseless kids? We are never sure what the issue is—is it retaliation? Jealousy, perhaps? Do they possess psychopath tendencies? Thus, the three main subheadings of emotions, control, and mental illness can be used to categorize criminal patterns.

Criminal behaviors refer to the behaviors of a lawbreaker which lead to the commission of an illicit act. In general, three categories—psychological, biological, and social—are involved in theories of criminal behavior. In actuality, human behavior is the result of interactions between numerous variables.

Human social communication relies heavily on facial expressions. Movement of the facial and skin-connected muscles of the face is what causes them. These muscles cause the skin to move, forming folds and lines that move face features like the mouth and eyebrows. The ability to perceive facial expressions is a crucial aspect of nonverbal communication, according to Anderson and Slotkin [1]. The truth is that we won't fully understand what someone is saying if we only pay attention to their spoken words while neglecting their facial expressions because words do not always accurately convey how they are feeling.

One of the main ways to identify someone, communicate with others, transmit information, and allude to emotions is through their face. We might be surprised by what our faces reveal. One's race, gender, health, emotions, psychology, and occupation can all be inferred from face features.

Criminal tendencies are one of the physical characteristics of humans, as such could be determined by facial expression. In order to infer facts about a person's personality, physiognomy is the study of a person's physical features, particularly their face. In the United States, facial recognition technology has enabled security officers to compare the faces of criminal suspects to images on social media, driver's licenses, and mug shots.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

A deep learning network architecture that learns directly from data is known as a convolutional neural network (CNN) or (CONVNET). CNN are notably helpful for identifying patterns in images to identify items, classes, and categories. CNN is a special type of deep learning network architecture that is employed for image identification tasks that require the processing of pixel input. Because of the convolutional layer's inherent ability to reduce high dimensionality of images without sacrificing information, CNN is particularly well suited for image processing. This paper sought to construct a machine learning model based on the FER2013 dataset using traditional CNN architecture that recognize the three basic categories of criminal behavior theories to determine criminal tendencies on suspected criminals.

Without the invention of machine learning (ML), the world would not exist as we know it. More benefits than at any previous time in history are provided to humans by its use in information processing. With the aid of machine learning, it is now possible to perform categorizations on a vast array of items and in contexts where humans would otherwise be unable to. But because it calls for particular theoretical underpinnings and the development of expertise to achieve the tuning of these algorithms, the development of ML algorithms is challenging. This paper focuses on the study of convolutional neural networks (CNN) and the creation of tests that make it easier to comprehend the impacts of their fundamental parameters. This is due to the broad variety of algorithms and methodologies available for performing element classification [2] [3].

The face is the most important feature for several functions, including identification, information transmission, interpersonal communication, and inferring emotions. Unexpected information may be revealed on our faces. The characteristics of a person's race, gender, age, health, emotion, psychology, and profession; can be revealed by their face [4] [5]. When analyzing facial photos to identify personality traits, machine learning has proven to be more successful than people [6]. Geng et al. [7] trained a machine to estimate age based on face scans. In order to identify depression and psychiatric disorders in Instagram facial photos, Reece and Danforth James et. al. [8] used a combination of machine learning algorithms and image processing.

Facial emotion recognition aims to teach a computer how to identify between six emotional emotions on the face: surprise, happiness, sorrow, disgust, rage, and fear [9]. Among the methods used for categorizing facial emotions are the fuzzy inference system developed by Pramerdorfer and Kampel [10], the hidden Markov model developed by Ituma et al [11] based on real-time tracking of the mouth shape, and the Bayesian network developed by Zhang [12].

Another personality feature is a propensity for crime. The first to mention that criminals may be recognized by their face features and emotions was Lombroso

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

[13] in 1871. Simonyan and Zisserman [14] recently examined this notion and quantified the relationship between criminality and face characteristics. They claimed that their computer can detect a criminal face with 90% accuracy after training four classifiers: logistic regression, k closest neighbors (KNN), support vector machines (SVM), and convolutional neural network (CNN).

Gender, race, and emotional expression on the face were all controlled in their model. The Convolutional Neural Network (CNN) has demonstrated significant potential in image processing, according to Ramos-Michel et al. [2]. A CNN is composed of convolutional, maximum pooling, activation, batch normalization, drop out, and fully linked layers. The training process is sped up by all of these layers. Different pooling techniques, including as average pooling and max pooling, are used to downsample the inputs and aid in generalization [2] [1]. Dropout, regularization, and data augmentation are employed to prevent overfitting. Using batch normalization, gradient vanishing and bursting was avoided. The development of several optimization methods used in training has also received a lot of attention [1] [10] [15] . Using the appropriate optimization technique can considerably improve a model's performance, despite the lack of a systematic theoretical manual for choosing an optimizer [16].

In terms of optimizer (SGD) usage, stochastic gradient descent is the most popular. It is a straightforward technique for altering a model's parameters in response to the gradient of a single data point. Adam combines the advantages of AdaGrad and RMSProp by changing the learning rate and incorporating gradient momentum [17] [18] [19]. The learning rate is an important factor in a CNN's training. A high learning rate could lead to loss divergence or oscillations at the minima. The model's convergence would be significantly slowed by a low learning rate, and it might even get stuck in an unfavorable local minimum. A common strategy is to utilize a learning rate scheduler that adjusts the learning rate as training progresses, with the learning rate decreasing either linearly or exponentially as the number of iterations increases [20] [21] [18] [22]. Step decay lowers the learning rate after a predetermined number of epochs by a factor. An adaptive learning rate schedule tries to automatically change the learning rate in accordance with local gradients during training. As a result of numerous [23], CNNs have developed into a helpful tool for image-related tasks [23].

Numerous developments have led to CNNs becoming a viable tool for image-related applications. In 2013, Ian Goodfellow et al. established the FER2013 database, which contains pictures with facial emotions, for a Kaggle competition on facial emotion detection [19]. It includes 35887 48x48 pixel 8-bit grayscale images of people with facial expressions that have been divided into three categories: 28709 training data, 3589 test data, and 3589 validation dataThe seven main categories of facial emotions—anger, disgust, fear, happiness, sadness, surprise, and neutral—have been applied to all of the database's photos. In order

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

to compare model performance in emotion recognition, the FER2013 dataset has been employed [24] [19] [25].

The classification accuracy of many CNN versions has been impressive, ranging from 70% to 76%. For instance, Pramerdorfer and Kampel [10] employed various CNNs and Ensemble them to increase the performance to obtain an accuracy of 75.8%, while Goodfellow et al. [23] used Ensemble ResMaskingNet of 6 CNNs to achieve an accuracy of 76.82%. Fusion CNN + Bag of Visual Words (BOVW) were utilized by Georgescu et al. [20] to reach 75.42% accuracy. Another ensemble of CNNs was employed by Pramerdorfer and Kampel [10] to reach an accuracy of 75.2% [10] [23] [26].

To attain high accuracy for the dataset, the researchers also used a variety of pre-trained models. When it comes to improving performance, assembling has shown exceptional performance. However, we plan to optimize the ensemble's constituent parts, a single network, to enhance ensemble performance even further. By including supplemental training data, other research has sought to enhance their model's performance on the FER2013; however, this is outside the purview of this work [2] [19].

Figure 1: Five Basic Facial Expression Emotional Features [2]

2. METHODS

2.1. Research Methods

At the course of the research, 6 convolutional layers, three dense layers and a batch size of 128 was used to construct a CNN model from a pretrained GoogLeNet network, which produced an initial accuracy test of 76%. Similarly, a smoothly decrease was observed in the trained losses over a tested length of the epoch while the test loss was constant at various fluctuations. These shows that when predicting the test data, the model is unable to generalize perfectly from the trained data.

Over fitting occurs accounting for noise in the trained data whenever the model attempts to learn too many details in the training data which makes the performances of the model on invisible or test datasets to be very poor. Because of these, the network cannot generalized the patterns found in the training datasets. Sequel to these over fittings, the accuracy of this prediction model on the

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

test dataset was raised to 75% and then fluctuates simultaneously at this point; whilst the accuracy further raised to 95% smoothly over the total length of the epochs on the train dataset.

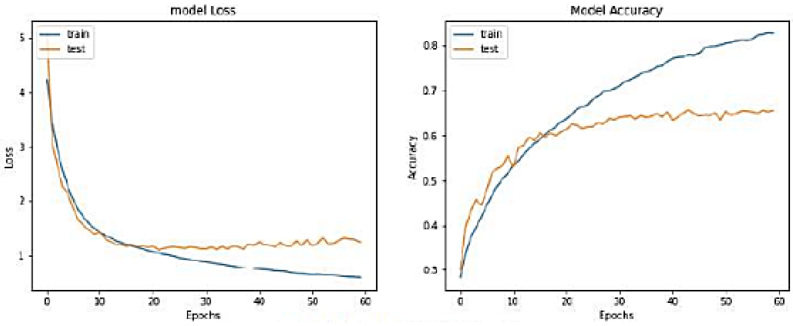

Figure 2. Model Accuracy Plot

To improve the accuracy through the applications of hyperparameter optimization, several steps were engaged. At first, from previous research, it was observed that the accuracy was obtained through the implementations of 6 convolution layers or 8 convolution layers respectively; this cause us to carry out hyperparameter optimization for 5 convolution layer model and 8 convolution layer model [18] [21] [27]. Secondly, through experimentations of difference batch sizes, notably 32, 64 and 128, it was found that 32 batch size yielded the best results. So the final batch size was settled to be 32. Finally, various optimizers were successively tried.

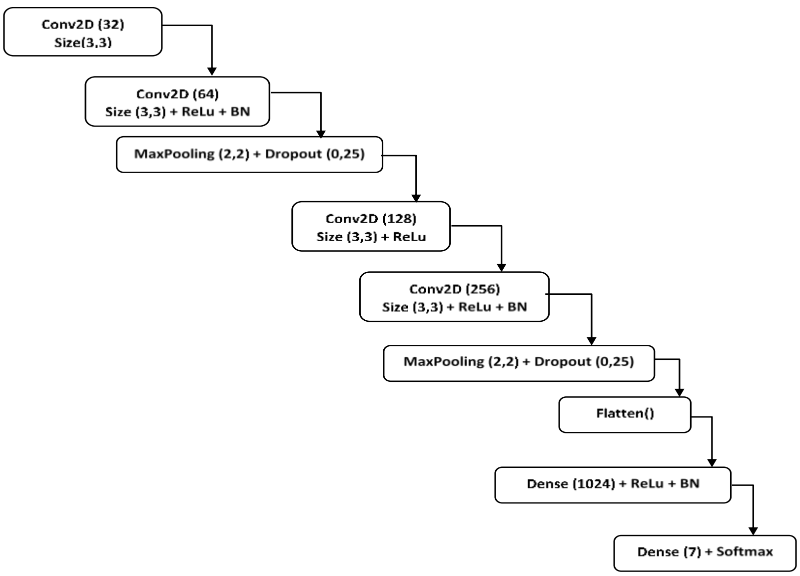

Figure 3. Proposed CCN CTP Architectural Model [14] [16]

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

2.2 The ConvNet Models

Convolutional neural networks are synthetic neural networks that, in addition to fully connected layers, also contain one or more convolutional layers.

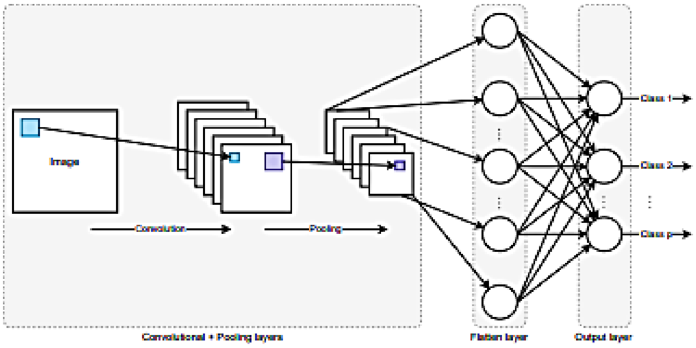

Figure 4. The CNN Architecture

Through the previously mentioned nodes and activation functions, it serves the purpose of processing the data obtained from the input layer. There is a direct connection between every node in the input layer and every node in the first hidden layer. The similar thing occurs between the first hidden layer's nodes and the second, and so on, up until the neural network's final layer. The output layer is the final layer. According to how the network has been structured, the output layer is in charge of doing the information's final treatment and delivering to the user the results of the data treatment. Convolutional layers are used to extract the key traits from the input tensor's data. The CNN's primary stages are depicted in Figure 6, along with a hypothetical and potential setup.

2.3 Mathematical Equations

2.3.1 Convolutional Layer

This layer convolves a kernel and the input data. Assume the kernel size of a single layer image is (3 × 3) pixels, such that it has a filter which poses an area of equal dimensions over the image. The convolution starts applying the Equation 5 in this position

$$conv = \sum_{i=1}^{r} \sum_{j=1}^{r} K_{ij} I_{ij} \qquad (1)$$

Where: I represent the kernel's elements, K describes the image area covered by the kernel

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

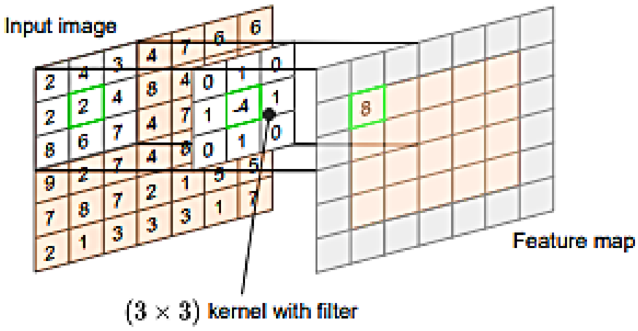

Figure 5. The Feature Map

The convolution finishes when the kernel passes over all the tensor positions. Figure 5 shows the creation of a new output tensor that is called the feature map, which contains the image's features by the kernel and image convolution.

2.3.2 Pooling Layer

The max pooling selected for each input and output elements could be expressed in matrix as swown in Equation 2 to 4.

$$vec\,y = S(x) vec\,x \tag{2}$$

Differentiating with respect to x, we have:

$$\frac{dz}{d(vec\,x)} = \frac{dz}{d(vec\,y)} S(x) \tag{3}$$

Hence.

$$vec\,y = S(x) vec\,x, \ \frac{dz}{d(vec\,x)} = \frac{dz}{d(vec\,y)} S(x) \tag{4}$$

Where r and c represent the number of rows and columns of both matrices.

2.3.3 Rectified Linear (ReLu)

The ReLU's introduced non-linearity in our ConvNet. Because the real-world data would want the network to learn with a non-negative linear value, hence, the need for ReLu Function. The ReLU function is expressed in Equation 5

$$vec\,y = diag\,S\ vec\,x, \ \frac{dz}{d\,vec\,x} = diag\,S \frac{dz}{d\,vec\,y} \tag{5}$$

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

2.3.4 Sigmoid Function

The sigmoid function is given as Equation 6.

$$\frac{dz}{dx} = \frac{dz}{dy} \circ y \circ \left(11^T - y\right) \quad (6)$$

With further differentiations, we have.

$$\frac{dz}{dx_{ij}k} = \frac{dz}{dy_{ij}d}\frac{dy_{ij}d}{dx_{ij}d} = \frac{dz}{dy_{ij}d}\frac{-1}{\left(1+e^{-x_{ij}d}\right)^2}\left(-e^{-x_{ij}d}\right) = \frac{dz}{dy_{ij}d}y_{ij}d\left(1-y_{ij}d\right)$$

2.3.5 Softmax

When dealing with categorization issues, the SoftMax function is an effective kind of activation. It is divided by the sum of the outputs and squeezed the outputs for each class between 0 and 1. The sofmax function is given as Equation 7.

$$\frac{e^{y_i}}{\sum_{i=1}^{k} e^{y_i}} \quad (7)$$

3. RESULTS AND DISCUSSION

3.1 Model Implementation

The dataset was captured from the FER2013 image dataset containing about 7070 images to five classes, these data set were classified using the six facial features as stated before which are the shape of the face, the eyebrows, eyeball, pupils, nostrils and lips. The model was implemented using spyder python 3.9 running on the anaconda integrated development environment (IDE) [28] [16] [18]. This programming editor was used to carry out the training of the image data sets at epoch 30, epoch 60, epoch 150 and epoch 300 respectively. A total of 7,837,895 parameters was used, where about 7,832,519 were training parameters and about 5,376 were testing parameters. The images were classified in batches using the features to determine the possibility of a person having criminal intention or not; figure 6 and figure 7 shows both the set of images as they were passed through various level of scaling as well as the scaling images using the proposed features examinations.

Figure 6. The sample images present various facial expressionexpressionsinal tendencies.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Figure 7. The scaled images presenting various facial expressionexpressionsinal tendencies using the features.

The group of images were classified into five categories representing the basic features as characterized by this research which help in predicting the motives of each person. In figure 7, these features are highlighted. The research considers the fact that, immaterial of the emotion of a person, such person may have a criminal tendency. That means, a person that emotionally shows to be 'Neutral' may commit crime or have the tendency to commit crime. Sequel to this, the features like the eyebrows, eyeball, pupils, nostrils and lips, becomes inevitable in deciding ones tendency. Upon the input images as depicted in figure 6, the system firstly carry out a preprocessing on these image, which helps in the eliminations of unwanted elements such as noise from the images; thereafter, the images are classified according to their landmarks that is, the distance between these features like that between the eyes, the length of the positioning of the nostrils etc. The search are run through the image database to find its perfect match and put them in groups to determine their mode. In other to enhance the optimization of hyperparameters, the Keras Tuner library was used. The library which contained Random Search helps in making combinations of parameters from a given search space to reach the best combination for the best performance. The following were the defined search spaces used:

a) A minimum of 32 and maximum of 256 number of kernels were used in the first convolutional layer with step size of 32.
b) A minimum of 64 and maximum of 512 number of kernels in other convolutional layer which are not the first layer with step size of 64
c) A minimum of 0 and maximum of 0.5 convolutional layer's dropout with step size of 0.05.
d) Finally, a minimum of 0 and maximum of 0.5 dense layer's dropout with step size of 0.05.

A total of six trials using Adam optimizer were carried out in five convolution layers as well as eight convolution layers at different epochs ranging from epoch 30 to epoch 300 and its validation accuracy noted. The output was determined using the activation functions that also allows for the normalization of each neuron's output to a range of 1 to 0 or -1 to 1. A ReLu Function with an output

Journal of Information Systems and Informatics
Vol. 6, No. 1, March 2024
p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

of $f(x) = \max(0, x)$ was introduced so as to introduce a non-linearity in our ConvNet.

```
def Create_CNN_Model():
    model = Sequential()
    #CNN1
    model.add(Conv2D(32, (3,3), activation='relu',
input_shape=(img_shape, img_shape, 3)))
    model.add(BatchNormalization())
    model.add(Conv2D(64,(3,3), activation='relu',
padding='same'))
    model.add(BatchNormalization())
    model.add(MaxPooling2D(pool_size=(2,2), padding='same'))
    model.add(Dropout(0.05))
    ;
;.    #CNN2
    model.add(Conv2D(256, (3,3), activation='relu', ))
    model.add(BatchNormalization())
    model.add(Conv2D(128,(3,3), activation='relu',
padding='same'))
    model.add(BatchNormalization())
    model.add(MaxPooling2D(pool_size=(2,2), padding='same'))
    model.add(Dropout(0.05))
```

Code 1. shows the model code in python

```
CNN_Predictions = CNN_Model.predict(test_data)
# Choosing highest probalbilty class in every prediction
CNN_Predictions = np.argmax(CNN_Predictions, axis=1)
test_data.class_indices
fig, ax= plt.subplots(figsize=(15,10))
cm=confusion_matrix(test_data.labels, CNN_Predictions)
sns.heatmap(cm, annot=True, fmt='g', ax=ax)
ax.set_xlabel('Predicted labels',fontsize=15,
fontweight='bold')
ax.set_ylabel('True labels', fontsize=15, fontweight='bold')
ax.set_title('CNN Confusion Matrix', fontsize=20,
fontweight='bold')
# Print classification report and confusion matrix
print('Classification report:')
print(classification_report(test_data.labels,
CNN_Predictions))
Criminal_Tendencies_Classes = ['shapes of the face '

                          'Eyebrows','Eyeball',
                                 'Pupils',
                                 'Nostrils',
```

Code 2. Shows the CNN prediction model code in python

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

3.2 Discussion

For the proposed model performance to be evaluated, Adam optimizer with a constant learning rate of 0.0001 was used on the targeted image dataset to train the model at epoch 300. With the dataset batch size of 32, these datasets was adjusted in form of zoom range of 0.1, rotation range of ±10o rotations width shift range and height ranges of 0.1 alongside of horizontal flip settings of 32.

In other to adequately validate the performance of the proposed model, sets of experiments were conducted using the targeted dataset collection and annotation making up of about 7070 images. As presented in Table 1, this model has achieved better performance than ordinary Neural Network which is what was expected due to the fact that it is made up of many layers. A pre-trained network was implemented and this pre-trained network learned about six (6) common features from the FER2013 image dataset and therefore retrained it on a smaller weight updates.

Table 1. Evaluations of the ConvNet model's results

Model	Training Accuracy	Validation Accuracy	Classifier Acurracy	Losses
ConvNet [FER 2013]	0.885	0.795	0.765	0.821
Pre-trained ConvNet	0.822	0.409	0.372	0.512
Adam Model	0.906	0.895	0.893	0.232

From table 1 above, `Training Accuracy of the ConvNet using the FER2013 image dataset is 85.5% while the Validation Accuracy and the Classifier Accuracy are 79.5% and 76.5% respectively, with a Total Losses of 82.1%. The Pre-trained ConvNet, has a Training Accuracy of 82.2% while the Validation Accuracy and the Classifier Accuracy are 40.9% and 37.2% respectively with a Total Losses of 51.2%. However, the proposed model has a Training Accuracy of 90.6 % while the Validation Accuracy and the Classifier Accuracy are 89.5 % and 89.3% respectively, with a Total Losses of 23.2%; making the proposed model a better option for the prediction of criminal tendencies in human being using their facial expressions. Issues bothering on the big differences that existed between the natural images in the FER2013 image datasets and the classified images in the grouped dataset causes the inability of the pre-trained model to provide better results.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

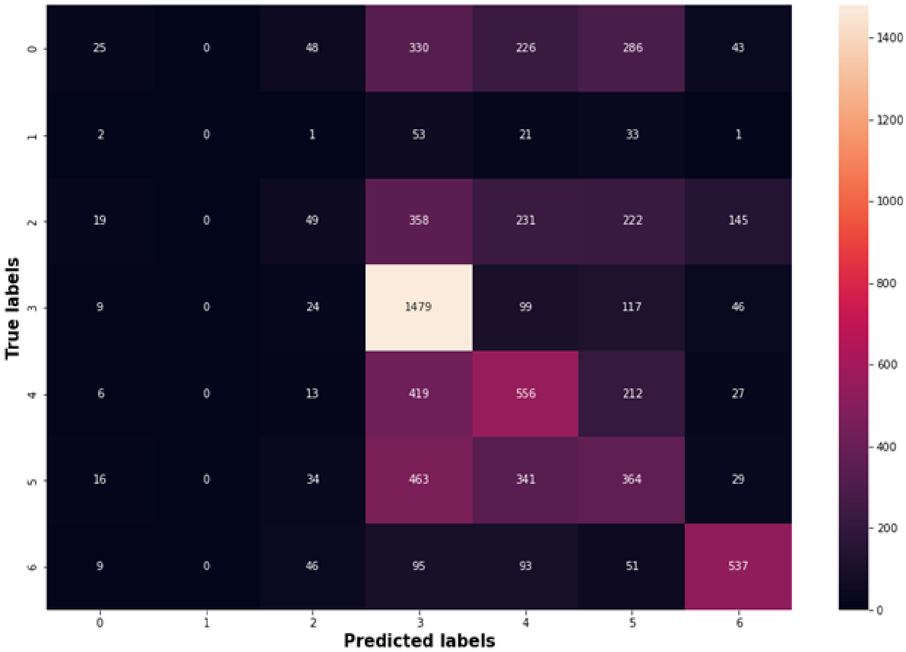

Figure 8: Convolutional Neural Network (CNN) model confusion matrix

The matrix shows plotting of the true labels against the predicted labels where the centroid is at ConvNet2D (3,3) which trained and validates a total of 1479 images. The output of a filter from the first convolutional layer and a filter from the second convolutional layer are displayed in Figure 9. The facial features that was used in the training process by the CNN and grouped within six (6) classes are highlighted in this figure.

The figure presents the resulting images of the evaluation and comparism of the facial features of people with criminal tendency and those without any criminal tendency. The said features are learned by the machine as a tools for classifications of two or more images in the training datasets. The highlighted features are the shapes of the face, the eyebrows, the top of the eye, pupils, nostrils, and lips. The features detected by the Max1(a, c, e) and Max2(b, d) convolutional layers in the deep learning model, representing someone with a criminal tendency Max12(a, b, d) against persons with no criminal tendency Max12(c, d, e) facial capturing.

Journal of Information Systems and Informatics
Vol. 6, No. 1, March 2024
p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

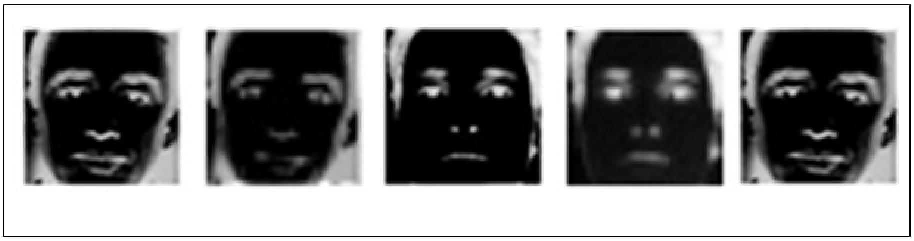

Figure 9. The Filter Outputs of the Convolutional Layers

Upon training of the convolution layer model from beginning at 300 epoch, the maximum accuracy of 90.6% was obtained. Even though the proposed model had a very high accuracy, a considerable level of over fittings was noticed which causes the inability of the system to be able to generalize the features adequately. Figure 10 shows the losses plots for this model while figure 13 shows the model's accuracy plots.

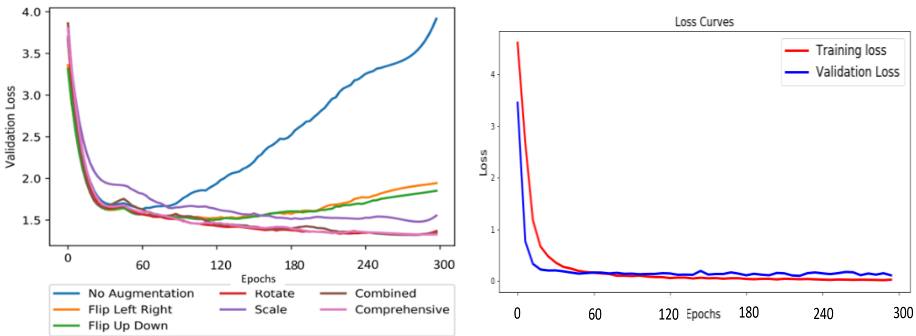

Figure 10. Model Loss Plots

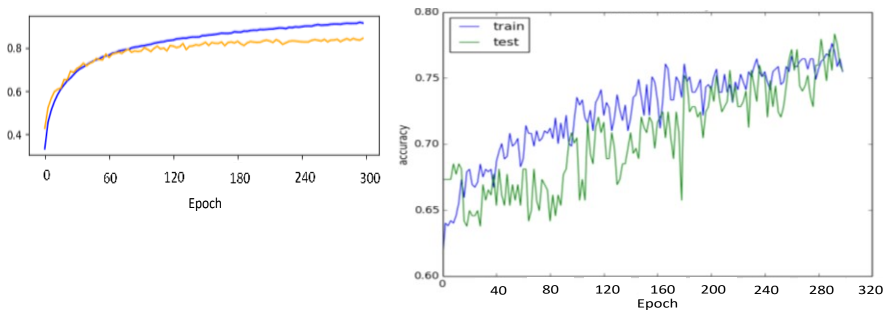

Figure 11. Model Accuracy Plots

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

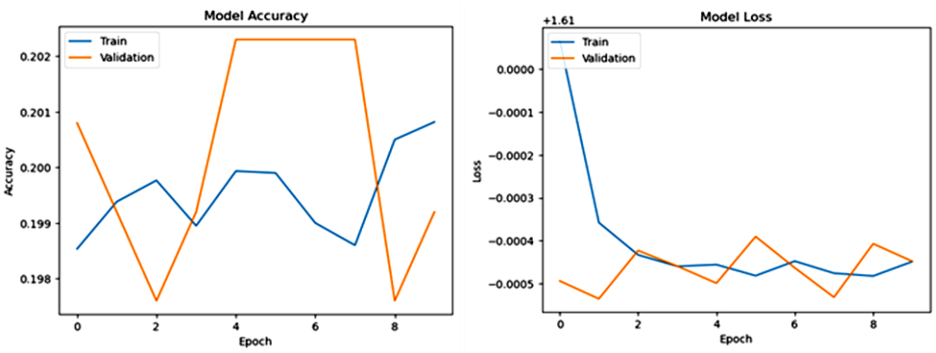

Figure 12. Model Loss vs. Accuracy Plots

4. CONCLUSION

The importance of images in semantic pattern recognition has been greatly increased thanks to the computing devices' explosive performance and memory space growth as well as the recent specialization of deep learning models [23], [15]. Facial photos may exhibit some personality attribute, similar to how a written post on social media displays the unique traits of its creator. In this paper, it was found out that the use of CNN in classifying images from the FER2013 image dataset to find the facial expressions of those with criminal intends and those without criminal intends was quite consistent in learning to reach its best test accuracy, which was 90.6% which means the use of CNN deep learning algorithm on features like the shape of the face, eyebrows, top of the eye, pupils, nostrils, and lips to classify the sets of images to those with criminal tendency and those without criminal tendency is effective and reliable. However, factors like emotion, age, gender, and background effects were observed to be major sources of impediment in effective classification of the facial images based on criminal and non-criminal tendencies. These factors where controlled through the elimination of non-neural facial images, elimination of images of children and elderlies, as well as cropping of the facial area out of the images.

REFERENCES

[1] T. R. Anderson and T. A. Slotkin, "Maturation of the adrenal medulla--IV. Effects of morphine," *Biochem. Pharmacol.*, vol. 24, no. 16, pp. 1469–1474, Aug. 1975, doi: 10.1016/0006-2952(75)90020-9.

[2] A. Ramos-Michel, M. Pérez-Cisneros, E. Cuevas, and D. Zaldivar, "Image Classification with Convolutional Neural Networks," in *Metaheuristics in Machine Learning: Theory and Applications*, vol. 967, D. Oliva, E. H. Houssein, and S. Hinojosa, Eds., in Studies in Computational Intelligence, vol. 967. ,

Cham: Springer International Publishing, 2021, pp. 445–473. doi: 10.1007/978-3-030-70542-8_18.
[3] A. J. Sarmah, N. Bhattacharyya, R. K. Jain, and P. Maroti, "Emotion Recognition Through Facial Expressions Using CNN," vol. 13, no. 2, 2023.
[4] M. Hashemi and M. Hall, "Retracted Article: Criminal tendency detection from facial images and the gender bias effect," *J. Big Data*, vol. 7, no. 1, p. 2, Dec. 2020, doi: 10.1186/s40537-019-0282-4.
[5] A. G. Reece and C. M. Danforth, "Instagram photos reveal predictive markers of depression," *EPJ Data Sci.*, vol. 6, no. 1, p. 15, Dec. 2017, doi: 10.1140/epjds/s13688-017-0110-z.
[6] S. Ruder, "An overview of gradient descent optimization algorithms." arXiv, Jun. 15, 2017. Accessed: Feb. 06, 2024.
[7] X. Geng, K. Smith-Miles, and Z.-H. Zhou, "Facial Age Estimation by Learning from Label Distributions".
[8] G.G. James, A.E. Okpako, and J.N. Ndunagu, "Fuzzy cluster means algorithm for the diagnosis of confusable disease," vol. 23, no. 1, Mar. 2017.
[9] P. Ekman and W. V. Friesen, "Facial Action Coding System." Jan. 14, 2019. doi: 10.1037/t27734-000.
[10] C. Pramerdorfer and M. Kampel, "Facial Expression Recognition using Convolutional Neural Networks: State of the Art." arXiv, Dec. 08, 2016. Accessed: Feb. 06, 2024.
[11] C. Ituma, G. G. James, and F. U. Onu, "Implementation of Intelligent Document Retrieval Model Using Neuro-Fuzzy Technology," *Int. J. Eng. Appl. Sci. Technol.*, vol. 4, no. 10, pp. 65–74, Feb. 2020, doi: 10.33564/ijeast.2020.v04i10.013.
[12] A. O. Vorontsov and A. N. Averkin, "Comparison of Different Convolution Neural Network Architectures for The Solution of The Problem of Emotion Recognition By Facial Expression," 2018.
[13] Cesare Lombroso, *Criminal Man*. Duke University Press, 2006.
[14] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition." arXiv, Apr. 10, 2015. Accessed: Feb. 06, 2024.
[15] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization." arXiv, Jan. 29, 2017. Accessed: Feb. 06, 2024.
[16] N. Tsapatsoulis, K. Karpouzis, G. Stamou, F. Piat, and S. Kollias, "A Fuzzy System for Emotion Classification Based on the MPEG-4 Facial Definition Parameter Set".
[17] N. Oliver, A. Pentland, and F. Bérard, "LAFTER: a real-time face and lips tracker with facial expression recognition," *Pattern Recognit.*, vol. 33, no. 8, pp. 1369–1382, Aug. 2000, doi: 10.1016/S0031-3203(99)00113-2.
[18] T. Connie, M. Al-Shabi, W. P. Cheah, and M. Goh, "Facial Expression Recognition Using a Hybrid CNN–SIFT Aggregator," in *Multi-disciplinary Trends in Artificial Intelligence*, vol. 10607, S. Phon-Amnuaisuk, S.-P. Ang, and S.-Y. Lee, Eds., in Lecture Notes in Computer Science, vol. 10607. , Cham:

Springer International Publishing, 2017, pp. 139–149. doi: 10.1007/978-3-319-69456-6_12.
[19] Chukwu, E. G., James; G. G; Benson-Emenike, M. E.; Michael, N. A, "Observed and Evaluated Service Quality on Patients Waiting Time of University of UYO Teaching Hospital using Queuing Models," *International Journal of Innovative Science and Research Technology*, vol. 8, no. 5, pp. 2094–2098, May 2023.
[20] M. I. Georgescu, R. T. Ionescu, and M. Popescu, "Local Learning with Deep and Handcrafted Features for Facial Expression Recognition," *IEEE Access*, vol. 7, pp. 64827–64836, May 2019.
[21] A. Khanzada, C. Bai, and F. T. Celepcikay, "Facial Expression Recognition with Deep Learning." arXiv, Apr. 07, 2020. Accessed: Feb. 06, 2024.
[22] B.-K. Kim, S.-Y. Dong, J. Roh, G. Kim, and S.-Y. Lee, "Fusing Aligned and Non-aligned Face Information for Automatic Affect Recognition in the Wild: A Deep Learning Approach," in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 1499–1508. doi: 10.1109/CVPRW.2016.187.
[23] I. J. Goodfellow *et al.*, "Challenges in Representation Learning: A report on three machine learning contests." arXiv, Jul. 01, 2013. Accessed: Feb. 06, 2024. [Online]. Available: http://arxiv.org/abs/1307.0414
[24] James, G. G.; Chukwu, E. G.; Ekwe, P. O.; ASOGWA, E. C; Darlington, C. H, "Design of an Intelligent based System for the Diagnosis of Lung Cancer," *International Journal of Innovative Science and Research Technology*, vol. 8, no. 6, pp. 791–796, Jun. 2023.
[25] R. T. Ionescu, M. Popescu, and C. Grozea, "Local Learning to Improve Bag of Visual Words Model for Facial Expression Recognition".
[26] Umoh, U. A., Umoh, A. A., James, G. G., Oton, U. U., Udoudo, J. J., B.Eng., "Design of Pattern Recognition System for the Diagnosis of Gonorrhea Disease," *International Journal of Scientific & Technology Research*, pp. 74–79, Jun. 2012.
[27] G. G. James, A. E. Okpako, C. Ituma, and J. E. Asuquo, "Development of Hybrid Intelligent based Information Retreival Technique," *Int. J. Comput. Appl.*, vol. 184, no. 34, pp. 1–13, Oct. 2022, doi: 10.5120/ijca2022922401.
[28] Y. Wang and M. Kosinski, "Deep Neural Networks Can Detect Sexual Orientation from Faces," Graduate School of Business, Stanford University, Stanford, CA94305, USA, USA, 2020.