iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52274

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Detection of Attention Deficit/Hyperactivity Disorder (ADHD)

Asst. Prof. Pavithra R[1], Anuu Karthiga I S[2], Jeyashree A[3], Sreemathi Sampath[4], Sruthi V[5]
[1, 2, 3, 4, 5]*Computer Science and Engineering, Coimbatore Institute of Technology, Tamil Nadu, India*

Abstract: ***In recent years, Machine Learning has had a huge impact on a range of technological disciplines. As technology expands, machine learning provides an exciting opportunity in health care to help medical professionals care for patients and manage clinical data. It improves the accuracy of diagnoses, personalize health care, and find novel solutions to decades-old problems. Hospitals and health care companies have begun to recognize the ability of machine learning to improve decision-making and reduce risk in the medical field, which has led to several new and exciting career opportunities. Attention Deficit/Hyperactivity Disorder (ADHD) is a chronic condition generally characterized by hyperactivity, impulsiveness, and difficulty in paying attention. The early diagnosis of ADHD is highly desirable, and there is a need for developing assistive tools to support the diagnosis process in this regard. At present, there are two main methods available - using fMRI data and tracking eye movements. The proposed approach offers a method to assist with the ADHD diagnosis with an emphasis on young children who are still in the early stages of development. It tries to discover the eye-tracking patterns of ADHD using machine learning. The fundamental concept is to convert eye-tracking scan-paths into a visual representation, thus the diagnosis may be viewed as an image classification task.***

Keywords: ADHD, eye movements, CNN, K-fold Cross Validation

I. INTRODUCTION

Attention Deficit/ Hyperactivity Disorder (ADHD) is a chronic illness that affects millions of children and often continues into adulthood. It manifests by a combination of persistent problems, such as difficulty sustaining attention, hyperactivity, and impulsive behavior. They struggle with low self-esteem, troubled relationships, and poor performance in school. These symptoms may be regarded normal and ignored by conventional social myths.

Often, the diagnosing procedure involves a set of cognitive tests that may require lengthy clinical evaluations. Also, the variability of symptoms makes it harder to determine whether someone has ADHD. Computer-aided technologies have become more prevalent in this area to guide through the examination and assessment process. Electroencephalography, eye-tracking, and magnetic resonance imaging (MRI) are a few technologies used.

Eye-tracking is the process of capturing, tracking, and measuring eye movements or the absolute point of gaze (POG), which refers to the point where the eye gaze is focused in the visual scene. The eye-tracking technology received particular attention in ADHD context since abnormalities of gaze, convergence insufficiency has been recognized as a major problem in ADHD.

II. LITERATURE SURVEY

The paper[1] presents an effective approach for ADHD identification at an early stage by using functional Magnetic Resonance Imaging data for the resting-state brain. The proposed methodology is based on seed correlation which computes the functional connectivity between seeds and all other voxels within the brain. The data are pre-processed using normalization, motion correction, slice timing correction, band-pass filtering, and the co-registration. The seed-based approach is used in this study as the feature extraction mechanism from the fMRI dataset. The classification is done using Convolution Neural Network based on extracted seed correlations from different Default Mode Network (DMN) regions. The proposed method using correlation on DMN regions has shown significant accuracies between 84% and 86% to be used with CNN for the identification of ADHD.

Compared to deep belief network approach, the proposed system shows high accuracy assisted by the usage of CNN. The highest sensitivity and specificity values are obtained from MPC region compared to other regions. The major benefits of this study include higher accuracy, sensitivity and specificity values compared to the previous studies. The ADHD identification process is only supported by a single measure of fMRI.

The paper[2] extracts the most prominent risk factors for children with ADHD and proposes a machine learning based approach to classify children as either having ADHD or healthy.

©IJRASET: All Rights are Reserved

It makes use of data extracted from 45,779 children aged 3–17 years. Since the class label is highly imbalanced, it adopts a combination of oversampling and under-sampling approaches to make a balanced class label. LR is used to extract the significant factors for children with ADHD based on p-values (<0.05).

It presents eight ML-based classifiers such as RF, NB, DT, XGBoost, KNN, MLP, SVM, and 1D CNN to predict children with ADHD. It concluded that RF-based classifier provided the highest classification accuracy of 85.5%, sensitivity of 84.4%, specificity of 86.4%, and an AUC of 0.94. This study illustrated that LR with RF-based system could provide excellent accuracy for classifying and predicting children with ADHD and can assist physicians in detecting and treating children with ADHD at an early stage.

The paper[3] presents a methodology based on deep learning that processes the video data of long-range computer graphics of children with ADHD. It explores the clinical significance of video long-range Electro-Encephalogram (EEG) in the diagnosis of children with ADHD. It constructs the model, mainly including fully connected neural network models and two-dimensional convolutional neural networks. Recognition method proposed has good recognition accuracy for the brain wave difference data of children with ADHD. The use of brain electrical signal data to classify and recognize ADHD diagnosis is the most accurate. The diagnosis method of data pre-processing and deep learning model analysis proposed is very feasible.

The paper[4] proposes a VR technology including head mounted display(HMD), game technology and sensors are used to create an interactive panoramic virtual classroom scenario in which a blackboard embedded with listening test, CPT test, executive test, and visual memory test specially designed for attention and executive functions is incorporated and in addition to this, it also develops a new assessment & diagnosis system based on children's performance, behaviour & reaction in the above-mentioned four tests through an enormous and systematic design of a battery of visual & auditory distractions of different intensity levels, durations, and sequence. The VR provides a comfortable and a reality like set up and thus capture the data based on the scenario we require. The demerit of this methodology is that is really time consuming because of all the tests involved and also the accuracy of the result is uncertain since the behaviour and reaction of the children cannot be watched as accurate as possible.

The paper[5] proposes a rule based system with a set of if-then rules for data classification. It uses oculomotor tasks for data collection and dimension reduction methods for data pre-processing. Data is classified using a variety of decision tree algorithms, such as J48, Random Tree, Random Forest, Linear Model Tree (LMT), REPTree, Hoeffding Tree, and Decision Stump, as well as classification ruling techniques like PART and JRip. It was determined that the chosen rule creation algorithms had a precision of 0.81 to 0.83 on performance measures, while the rules generated by the classification ruling methods had a precision of 0.89, making them a better set of rules for the given dataset.

The paper[6] proposes the idea of detecting Attention deficit/hyperactivity disorder in a recorded eye movements environment together with information about the video stimulus in a free-viewing task. A deep-learning based sequence model is pre trained for the detection and convoluted neural networks (CNN) is applied which preprocess the raw eye-tracking which consists of horizontal and vertical sample data. The model is pre-trained with CNN and SWAN methodology and later it is fine tuned. The task-free nature of the viewing setting allows us to interpret eye movements to reflect differences in visual attention allocation between individuals with and without ADHD and it can be applied to any inference task that uses eye movements and a static or moving visual stimulus as input to detect ADHD.

III. PROPOSED SYSTEM

Recent work integrating an eye tracker with CPTs was proposed to be a feasible way of enhancing diagnostic precision and a combination of eye movement fixation and saccade features have been shown to be predictors of ADHD diagnosis in adults. The system using eye tracking technology can greatly reduce the effects of the disease when treated properly. In this regard, a prediction system using eye tracking technology is proposed which can be used as an assistive tool by the doctors to diagnose ADHD. The eye movements of the patients are recorded and are converted into scan paths using visualization techniques of eye-trackers. From these scan paths, dataset is obtained.

The dataset consists of eye visualizations of both the ADHD diagnosed patients and normal people in the age group between 5 to 8 years.

The data is pre-processed and the images are augmented to avoid overfitting, followed by dimensionality reduction. Then, CNN model is built by feeding it with data using k-fold optimization technique. Sigmoid activation function is used as it is suitable for binary classification.

The metrics of the model like accuracy, precision, etc. is checked and confusion matrix is displayed. The system overall can be viewed as an image classification system which predicts ADHD diagnosed patients.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

IV. METHODOLOGY

A. Eye Tracking

Eye-tracking is the process of capturing, tracking and measuring eye movements or the absolute point of gaze (POG), which refers to the point where the eye gaze is focused in the visual scene. Visual representations of eye-tracking scan paths can discriminate the gaze behaviour of ADHD. At its core, the key idea is to compactly render eye movements into an image-based format while maintaining the dynamic characteristics of eye motion (e.g. velocity, acceleration) using color gradients. In this manner, the prediction problem can be approached as an image classification task.

The SMI remote eye-tracker is used as the main instrument used to perform the eye-tracking function. The device belongs to the category of screen-based eye-trackers. It can be conveniently placed at the bottom of the screen of a desktop PC or laptop. Three basic categories of eye movements are aimed to be captured by eye-trackers including: i) Fixation, ii) Saccade, and iii) Blink.

A fixation is the brief moment that happens while pausing the gaze on an object in order that the brain can perform the perception process. The average duration of fixation was estimated to be 330 milliseconds. The accurate perception requires constant scanning of the object with rapid eye movements, which are so-called saccades. Saccades include quick, ballistic jumps of 20 or longer that take about 30–120 milliseconds each. On the other hand, a blink would often be a sign that the system has lost track of the eye gaze.

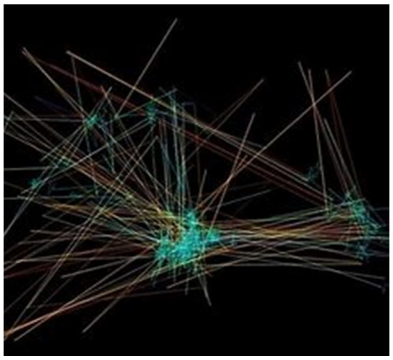


Fig 1. Visualisation of scan path of eye movement

The scan path visualizations are produced to represent the spatial coordinates of the eye movement along with its dynamics. These coordinates represents the change in participant's gaze direction into the screen with respect to time. Based on the change in position along associated time, the velocity of gaze movement can be calculated. The acceleration of movement could be computed based on the change in velocity, and the jerk of movement could be accordingly computed based on the change in acceleration. The change in colour across the line represented the movement dynamics. The values of RGB components were tuned based on velocity, acceleration and jerk with respect to time. Finally, the images are produced for further analysis.

B. Convolutional Neural Network (CNN)

A convolutional neural network is a feed-forward neural network that is generally used to analyse visual images by processing data with grid-like topology. It's also known as a ConvNet. A convolutional neural network is used to detect and classify objects in an image. A convolution neural network has multiple hidden layers that help in extracting information from an image. The four important layers in CNN are:

1) Convolution Layer

Convolution Layer is the first step in the process of extracting valuable features from an image. A convolution layer has several filters that perform the convolution operation. Every image is considered as a matrix of pixel values. The convolved feature matrix is obtained on computing the dot product by sliding the filter matrix/kernel over the input image. The convolved features are passed on to the next layer.

2) ReLU Layer

ReLU stands for the rectified linear unit. Once the feature maps are extracted, the next step is to move them to a ReLU layer. This layer performs an element-wise operation and sets all the negative pixels to 0. It introduces non-linearity to the network, and the generated output is a rectified feature map. FORMULA: R(z)=max(0,z)

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

3) Pooling Layer

Pooling is a down-sampling operation that reduces the dimensionality of the feature map. The rectified feature map now goes through a pooling layer to generate a pooled feature map. There are various types of pooling and above is the illustration of max pooling where the maximum of the all the values is taken into consideration. The next step in the process is called flattening. Flattening is used to convert all the resultant 2-Dimensional arrays from pooled feature maps into a single long continuous linear vector.

4) Fully Connected Layer

A fully connected layer refers to a neural network in which each neuron applies a linear transformation to the input vector through a weights matrix. As a result, all possible connections layer-to-layer are present and thus every input of the input vector influences every output of the output vector.

C. Principal Component Analysis(PCA)

The Principal Component Analysis is a popular unsupervised learning technique for reducing the dimensionality of data and at the same time increasing interpretability of data. It also minimizes information loss. It helps to find the most significant features in a dataset and makes the data easy for plotting in 2D and 3D. PCA helps in finding a sequence of linear combinations of variables. PCA is one of the most popular techniques for dimensionality reduction that has been widely applied in problems dealing with data of high dimensionality such as image compression and face recognition.

V. IMPLEMENTATION

The system is designed for the purpose of detecting ADHD in humans. It uses eye-tracking technology as abnormalities of gaze have been consistently recognized as the hallmark of ADHD. A system which does image classification is built using CNN which can predict two categories of participants (ADHD positive and negative). The system can be used as an assistive tool to support the diagnosis process.

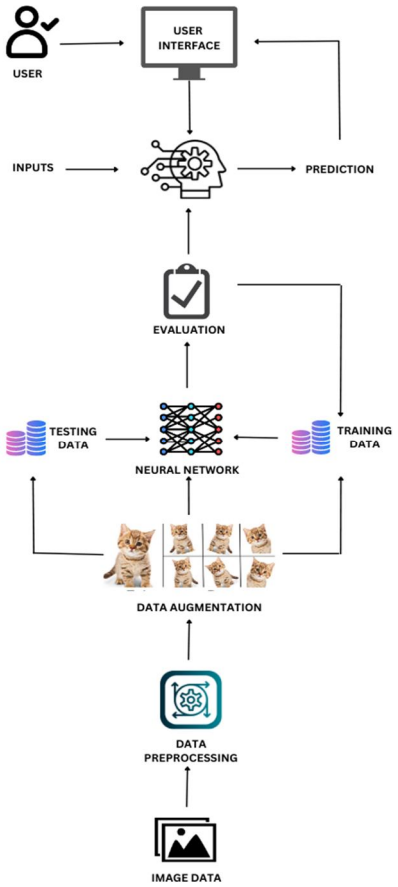


Fig 2. Architecture of the Proposed System

A. Data Acquisition

The dataset comes from a publicly accessible repository – Fig share. It contains images of eye - tracking patterns of ADHD positive and ADHD negative individuals recorded using SMI remote eye-tracker.

Three categories of eye movements are captured – Fixation, Saccade and Blink. The data is a scan path representing consecutive fixations and saccades as a trace through time. The change in color across the line represents movement dynamics. The line's color shift indicates the dynamics of movement. With relation to time, RGB components reflect velocity, acceleration, and jerk. Saccades are shown in the photos as yellow or white, with white denoting swift movements, whereas fixations are shown as cyan.

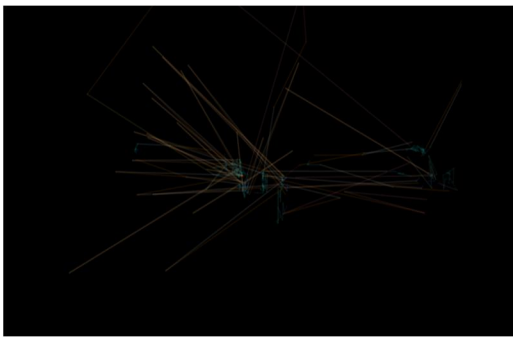


Fig 3. ADHD Positive Sample

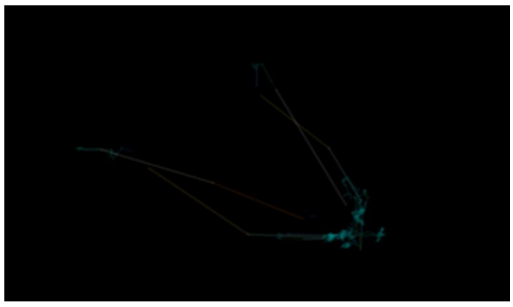


Fig 4. ADHD Negative Sample

B. Data Pre-Processing

Dimensionality refers to number of variables under consideration. As illustrated in Fig 5, the initial dimensionality was 640*480*3 (image size * RGB components) which gives more than 900K features which complicates the model training. Since most images had a lot of empty space and unused pixels, the images are first scaled down to 100*100 size to make things simpler. After that, they are converted to grayscale. Then Principal Component Analysis (PCA) is implemented to transform grayscale images into more compressed format. Using orthogonal transformations PCA converts a correlated set of data into linear uncorrelated set.

C. Image Augmentation

A notable method for enlarging datasets is image augmentation, which also improves model generalization and lowers the risk of overfitting. The fundamental goal of augmentation is to generate artificial samples by applying a wide range of picture transformations (e.g. rotation, shearing). It has been established that augmentation increases the predictability of image classification. Hence, augmentation is used to create different eye-scan path visualizations. An extra 2,735 samples were added to the collection, and for each visualization, five augmented images were created.

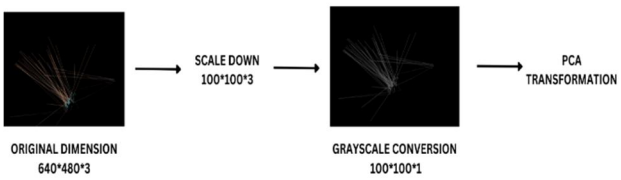


Fig 5. Stages of Data Pre-processing

D. Model Building

In machine learning, building a model is creating a mathematical representation by generalizing and learning from training data. The created machine learning model is then used to analyze new data and derive predictions.

E. Adding CNN Layers

CNN instantiated as a sequential model works in two steps: feature extraction and classification. Several filters and layers are applied to the images during feature extraction in order to extract information and features. ReLU and sigmoid activation functions are employed. The application of a pooling layer minimizes the size of the feature map and speeds up calculation.

F. Compiling The Model

Model compilation is a step that comes after writing the model's statements but before training really begins. It also defines the metrics, optimizer or learning rate, loss function, and checks for format issues. A compiled model is required for training but not for making predictions.

G. Training The Model

To detect patterns and generate predictions, the machine learning model is fed with the prepared data during training. It results in the model learning from the data so that it can accomplish the task set. The model improves in prediction over time with training.

```
model.summary()

Model: "sequential_9"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_36 (Conv2D)          (None, 256, 256, 16)      160

 max_pooling2d_36 (MaxPoolin  (None, 128, 128, 16)     0
 g2D)

 dropout_36 (Dropout)        (None, 128, 128, 16)      0

 conv2d_37 (Conv2D)          (None, 128, 128, 16)      2320

 max_pooling2d_37 (MaxPoolin  (None, 64, 64, 16)       0
 g2D)

 dropout_37 (Dropout)        (None, 64, 64, 16)        0

 conv2d_38 (Conv2D)          (None, 64, 64, 32)        4640

 max_pooling2d_38 (MaxPoolin  (None, 32, 32, 32)       0
 g2D)

 dropout_38 (Dropout)        (None, 32, 32, 32)        0

 conv2d_39 (Conv2D)          (None, 32, 32, 32)        9248

 max_pooling2d_39 (MaxPoolin  (None, 16, 16, 32)       0
 g2D)

 dropout_39 (Dropout)        (None, 16, 16, 32)        0

 flatten_9 (Flatten)         (None, 8192)              0

 dense_18 (Dense)            (None, 256)               2097408

 dense_19 (Dense)            (None, 1)                 257

=================================================================
Total params: 2,114,033
Trainable params: 2,114,033
Non-trainable params: 0
```

Fig 5. Summary of the model

H. Evaluation

Then the model's performance is checked by testing the performance of the model on previously unseen data. The unseen data used is the testing set which is split earlier. When used on testing data, we can get an accurate measure of how the model will perform and its speed. Performance metrics such as recall, precision, accuracy, ROC, AUC are displayed where ROC is a probability curve and AUC represents the degree or measure of separability.

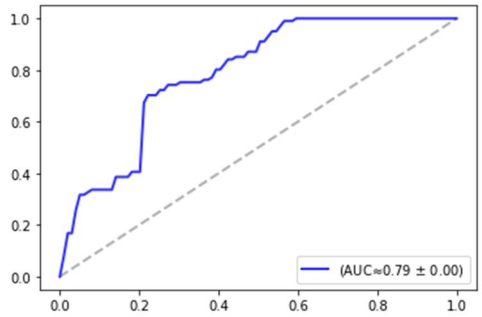


Fig 7. ROC Curve

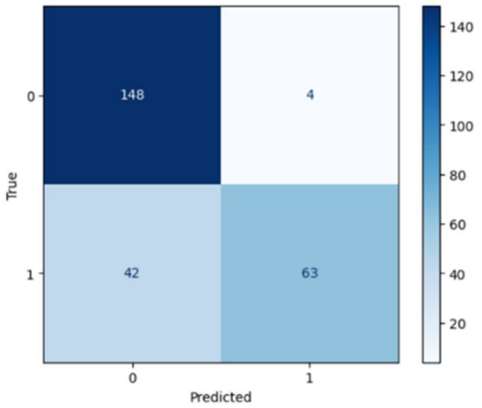


Fig 8. Confusion Matrix

VI. RESULTS AND DISCUSSION

The proposed system takes as input the eye-tracking visualization path and predicts whether it is the scan path of the ADHD positive or negative subject. Value of 0 represents ADHD negative sample and 1 represents ADHD positive sample. The drawback of the system is only a few eye-tracking data is available which reduces the accuracy of the model. Yet, optimization techniques such as k-fold cross validation is applied rather than splitting the data into train/test to improve the metrics of the model. 10-fold cross validation is used and the metrics are displayed. Data with augmentation and without augmentation is passed into the model and the metrics are compared. It is clear that model with augmented data has high performance score and improved metrics.

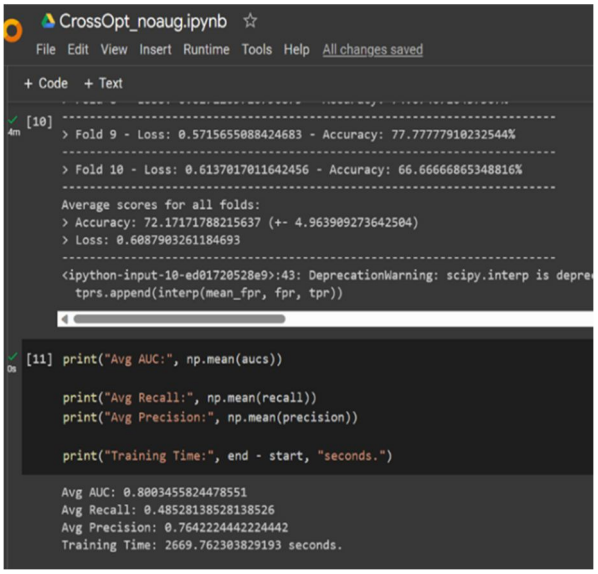


Fig 9. Metrics of the model without augmentation

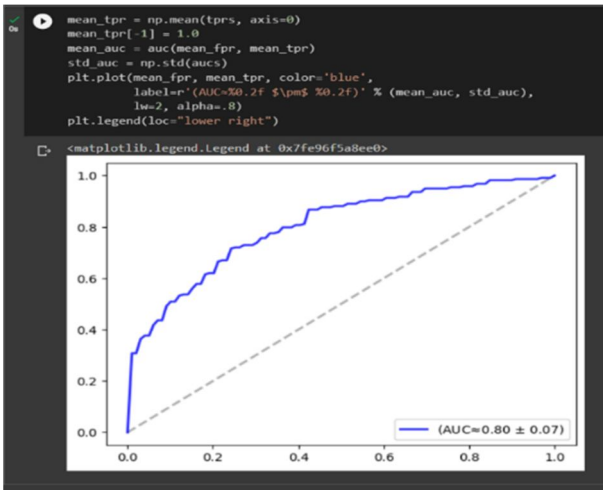


Fig 10. ROC Curve without augmentation

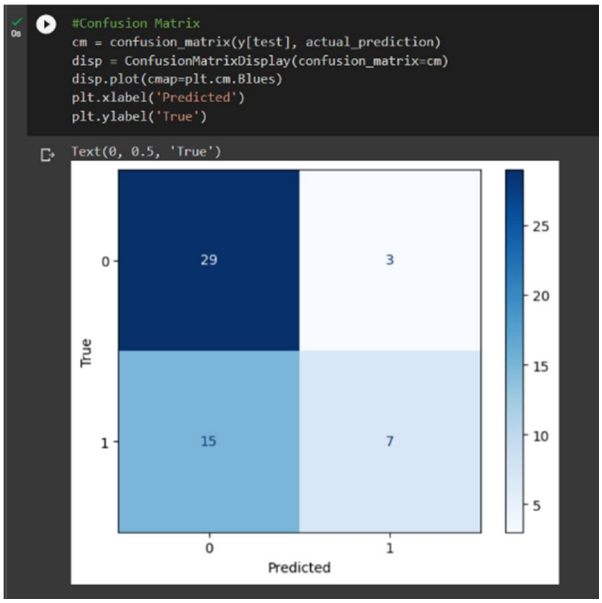


Fig 11. Confusion Matrix without augmentation

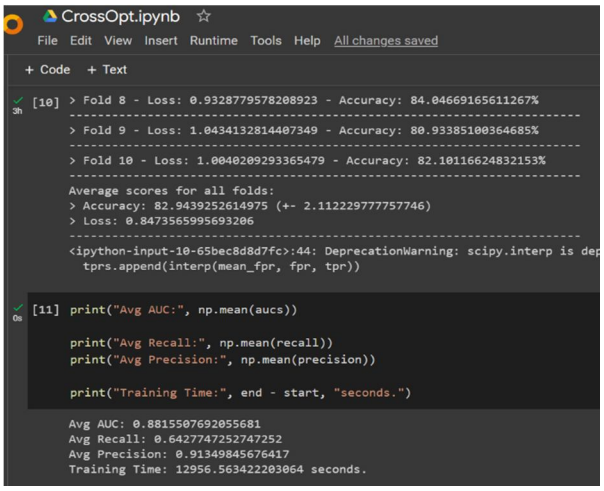


Fig 12. Metrics of the model with augmentation

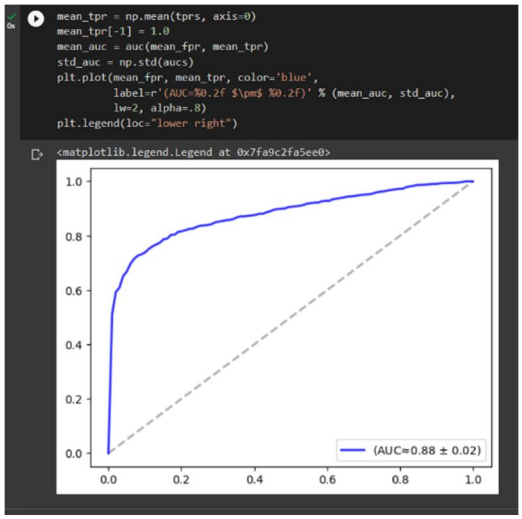


Fig 13. ROC Curve with augmentation

Fig 14. Confusion Matrix with augmentation

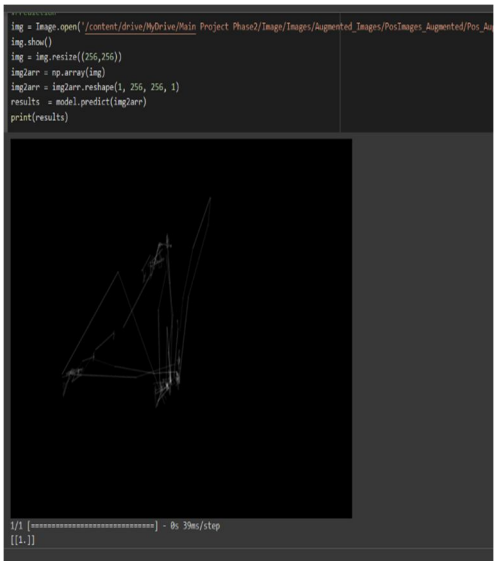


Fig 15. Prediction of ADHD Positive

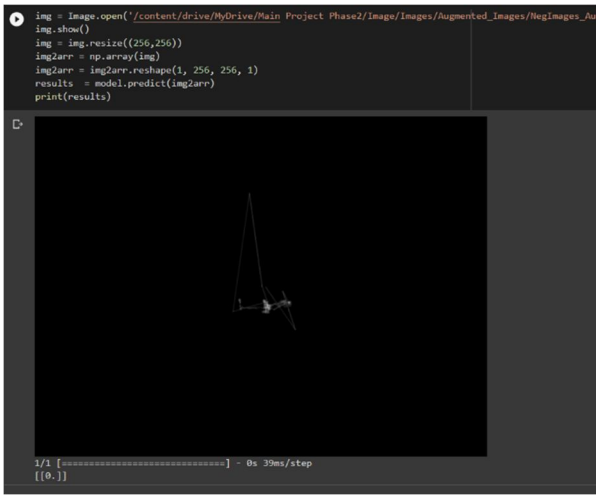


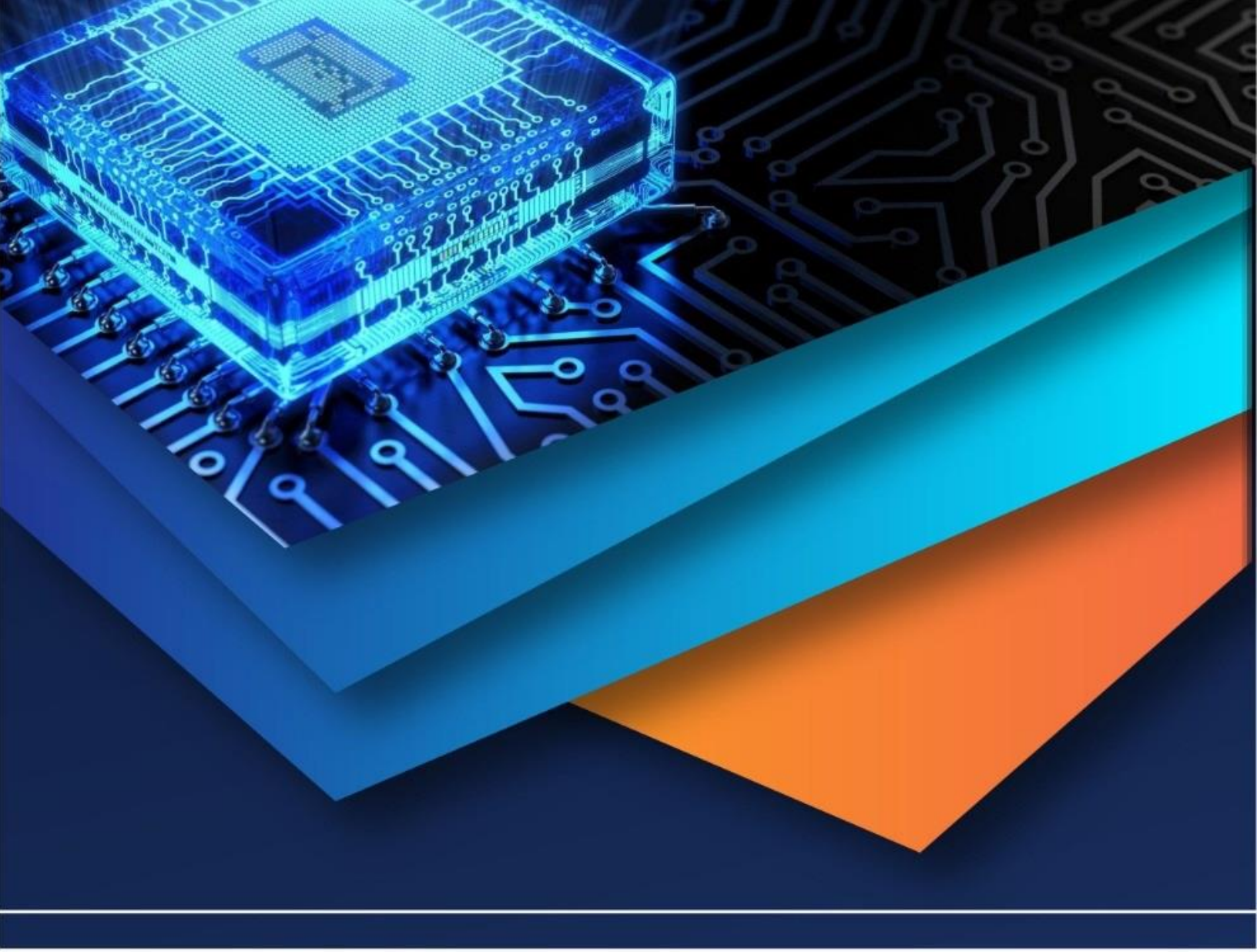
Fig 16. Prediction of ADHD Negative

VII. CONCLUSION

The aim of the project is to design a system which predicts ADHD positive and negative patients. The model is created using k-fold cross validation and CNN with a precision of 91.3% The work can be extended where other disorders like ASD can also be predicted using the system. The system should be developed in such a way that when the eye tracking image is uploaded, the system predicts whether the patient has ADHD, ASD or other neuro related disorders which can be predicted using eye-tracking data.

REFERENCES

[1] Gangani Ariyarathne., et al "ADHD Identification using Convolutional Neural Network with Seed-based Approach for fMRI data" – ICSCA Conference, 2020
[2] Md.Maniruzzaman., et al "Predicting children with ADHD using Behavioural Activity : A machine learning analysis" – Applied Sciences 12(5):2737, 2022
[3] Dingfu Zhou., et al "Deep Learning enabled diagnosis of children's ADHD based on big data of video screen long range EEG" – Hindawi, Journal of Healthcare Engineering, 2022
[4] Shih-Ching Yeh., et al "Innovative ADHD Assessment System using Virtual Reality" – IEEE EMBS Conference, 2012
[5] Senuri De Silva,. et al "A rule based system for ADHD identification using eye movement data" – Moratuwa Engineering Research Conference (MERCon), 2019
[6] Shuwen Deng,. et al "Detection of ADHD based on eye movements during natural viewing" - German Federal Ministry of Education and Research, 2022



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)