<u>Eye Gaze Tracking Using WebCam</u>

(Applied on Autism detection)

7th SEMESTER END-SEM MINI PROJECT REPORT

FOR THE DEGREE OF

BACHELOR OF TECHNOLOGY
IN

INFORMATION TECHNOLOGY

BY

AnkitaNasipuri (LIT2015012)
Puja Kumari (LIT2015017)
Samriddhi Niranjan (LIT2015021)
Kajol (LIT2015022)

UNDER THE SUPERVISION OF

Prof. AnupamAgrawal
IIIT-ALLAHABAD

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, LUCKNOW

(An Institute Established Under Public Private Partnership (PPP) Model By MHRD, UP Govt. & UPLC)

September, 2018

DECLARATION BY THE CANDIDATES

We hereby declare that the work presented in this mid semester project report of B.Tech (IT) 7th Semester entitled **"Eye Gaze Tracking Using WebCam (Applied on Autism Detection)"** submitted by us at Indian Institute of Information Technology, Lucknow, is an authenticated record of our original work carried out from August 2018 to November 2018 under the guidance of **Prof. AnupamAgrawal**.

Due acknowledgements have been made in the text to all other material used. The project was done in full compliance with the requirements and constraints of the prescribed curriculum.

Place: Allahabad

Date: 25 November, 2018

AnkitaNasipuri (LIT2015012)
Puja Kumari (LIT2015017)
Samriddhi Niranjan (LIT2015021)
Kajol (LIT2015022)

CERTIFICATE FROM SUPERVISOR

I do hereby recommend that the mid semester mini project report prepared under the supervision by AnkitaNasipuri (LIT2015012), Puja Kumari (LIT2015017), Samriddhi Niranjan (LIT2015021) and Kajol (LIT2015022) titled "Eye Gaze Tracking Using WebCam(applied on Autism Detection) " be accepted in the partial fulfillment of the requirements of VII semester of Bachelor of Technology in Information Technology for Examination.

Date: Prof. AnupamAgrawal
Place: Allahabad IIIT-Allahabad

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to the Department of Information Technology, IIIT Allahabad for providing valuable help, discussion, suggestions and resources. We are especially thankful to **Prof. Anupam Agrawal** for his valuable support and helping us figure out approach to the problem.

AnkitaNasipuri
Puja Kumari
Samriddhi Niranjan
Kajol

Table of Contents

Page no.

1. Introduction……………………………….…...………………...…………...1
2. Motivation….…………………………………….…....……………….2
3. Problem Objective and Scope……………...…………..........................….3
4. Literature Review ……………………………..........……………………...4
5. Methodology…………………………..6
6. Result……………………………….....…………....……...………...…...11
7. Software and Hardware Requirements …………………………………...13
8. Activity Time Chart……………………………………………………..…...14
9. Conclusion and Future Scope.……………………………………..…...15
10. References….……………………………………………………………..16

ABSTRACT

Autism Spectrum Disorder (ASD) is broadly definedas a pervasive developmental disorder characterized by a triad of impairments including social communication problems, difficulties with reciprocal social interactions, and unusual patterns of repetitive behaviour. The diagnosis of ASD remains a challenging task, which requires a set of cognitive tests and perhaps hours of clinical examinations.One of the characteristic hallmarks of ASD is the difficulty of making or maintaining eye contact. In this respect, the eye-tracking technology has come into prominence to support the study and analysis of autism. This project explores machine learning as a tool for autism detection using eye tracking. We present a comparative analysis of three approaches: Convolutional Neural Network (CNN), Deep Neural Network (DNN), and Non-Neural Network models (SVM and Random Forest Classifier).

1. INTRODUCTION

Autism Spectrum Disorder(ASD) is a lifelong condition generally characterized by social and communication impairments. One of the characteristics hallmarks of ASD is the difficulty of making or maintaining eye contact. ASD has been considered to affect about 1% of the world's population. Most common symptom of ASD is an impairment of ability to make and maintain eye contact. The study of eye gaze movements help to analyze human processing of visual information for interactive and diagnostic applications. It also provides feedback on when, where and for how long a respondent looks at a particular stimulus. Eye tracking has been used to investigate gaze behavior in normal population[2]. It is applied as an investigative method to obtain knowledge of significant characteristics of human behavior. Unlike people with typical development, children with autism do not look at people with eyes directed at them and cannot visually focus for a long time[1].

In this project we intend to build a tool to detect autism at a preclinical stage so that proper intervention can be provided for development of autistic children.We have done a comparative analysis of three machine learning approaches of autism detection using eye tracking, which are, CNN, DNN, and non-neural network methods – SVM and Random Forest Classifier.

For training we have used an image dataset based on a set of eye tracking experiments. The images are the visual gaze patterns of ASD diagnosed and non-ASD participants.

2. MOTIVATION

Autism, also known as the Autism Spectrum Disorder (ASD), is one of 5 Pervasive Development Disorder which causes a person to have severe difficulties in the areas of cognitive and social development. Autism falls under a category of neurological development disorder and brings disabilities in communication, learning, development and social skills.Prevalence of autism is 1 in every 88 people. It is a permanent problem which affects a high percentage of population. Difficulty of early detection of information include lack of information, inadequacy of development assessment tools and cost expensive initial diagnosis. Children who are early diagnosed with autism can have improved long-term prognosis. The project aims to propose a method for early diagnosis of autism which is efficient and inexpensive.

3. PROBLEM OBJECTIVE

To develop a software system which takes visual gaze patterns as input and predicts whether it belongs to autistic patient or not using three different approaches and compare them. The output given by the system will be autistic or non-autistic.

4.LITERATURE REVIEW

Sl. No.	Title	Conference Name/ Publisher	Year	Methodology Used	Merits	Demerits	Future Scope
1.	Diagnosis of autism using eye tracking system*	Global Humanitarian Technology Conference, IEEE	2016	Shows two videos simultaneously to the children and makes ocular edge detection for correctly identifying iris and quantifying black pixels of the iris to discriminate the direction of gaze to find autism in children.	Cost effective portable system, computation ally efficient.	Need to take into account contribution percentage of distractions in the result.	Distracti on of children can be handled and distractio n percentag e can be included in the result.
2.	Computer Vision Based Gaze Tracking for Accident Prevention	World Conference on Futuristic Trends in Research and Innovation for Social Welfare, IEEE	2016	Consists of data acquisition, division into frames, face tracking and feature extraction. Uses LLS algorithm and OECM for iris detection. Blink detection distinguishes voluntary and involuntary blinks.	Able to detect at day and night, does not require calibration. No major recalibration necessary if camera position is changed.	Monocular camera needs to be installed in the steering wheel.	
3.	Eye Gaze Tracking to Operate android-based communication helper application	International Electronics Symposium on Knowledge Creation and Intelligent Computing, IEEE	2017	The change in iris movement can be used to activate menu. Steps involve initial setup, eye localization, segmentation, eye center detection, pointer movement and menu click.	Success rate of left and right gaze tracking is 80% and straight gaze tracking is 90%.	Has many variables, problems occur when user makes some movement or doesn’t look at screen.	Can be extended to other applicati ons to aid for people with motoric disabiliti es.
4.	The application of Eye Tracking Technology in the study of Autism	The Physiological Society	2007	Illuminate eye with infrared beam. Identify reflection from cornea as brightest and pupil as second brightest image.	Allows direct, objective and quantitative analysis for autism detection.	In some cases no difference was found in eye gaze due to different participant groups.	Can be extended for behavior al and neuroima ging studies.

5.	Eye gaze tracking	International Conference on Computing and Communication Technologies, IEEE	2009	Consists of eye detection using rapid object detection, eye tracking using Lucas Kanade and detecting eye gaze using Gaussian process.	Uses Haar features and Cholesky decomposition which is faster and more stable.	Has to stable the position of head with the camera, if head moves after training, prediction may be wrong.	Can calculate the relation of movement of eye with head.
6.	Subpixel Eye Gaze Tracking	International Conference on Automatic Face and Gesture Recognition, IEEE	2002	Uses local pattern to achieve gaze estimation. Defines eye corner, perform 2D convolution and bicubic interpolation to get results.	Computationally efficient, lighting invariant.		Can be applied with head pose estimator Can enhance accuracy of head estimation.
7.	Real-Time Eye Gaze Tracking With an Unmodified Commodity Webcam Employing a Neural Network	CHI (ACM Conference on Human Factors in Computing Systems), ACM	2010	Employs ANN to estimate user's gaze. Face and eye detection, iris and pupil detection is done. Cropped eye sent to neural network where gaze coordinate is estimated.	Works on unmodified webcam. Low error (< 3.68°).	Sensitive to head movement. Longer training required to increase accuracy.	Can reduce user training time by pre training with large dataset.
8.	Detecting Gaze Towards Eyes in Natural Social Interactions and Its Use in Child Assessment	Proceedings on Interactive, Mobile, Wearable and Ubiquitous Technologies, ACM	2017	Adult wears point of view camera. Face detected, head pose estimated, eye features extracted, output predicted using random forest classifier.	Helps understand atypical gaze behavior and screening of autism.		Can be extended for analysis of social roles in group and turn taking.
9.	Eye Tracking for Everyone*	Conference on Computer Vision and Pattern Recognition, IEEE	2016	Train a CNN learned end to end for gaze prediction without the need to include manually engineered features like head pose.	Introduces a large scale dataset for mobile eye tracking. Allows CNN to make accurate predictions.		Can be used for various eye tracking applications on mobile devices.

** denotes base paper.*

5. METHODOLOGY

In this project, we have done a comparative analysis of three approaches of autism detection using eye tracking. The first approach involves training of a Convolutional Neural Network (CNN) to detect autism using eye gaze patterns. The second approach is the same as first approach but it involves training of a Deep Neural Network (DNN) to detect autism using gaze patterns. The third approach involves two Non-Neural Network methods (SVM and Random Forest Classifier) for autism detection.

The basic approach of autism detection consists of four steps. The four basic steps are Eye Gaze Tracking, Visual Pattern Generation, Data Augmentation and Classification as shown in figure 5.1. First, a video is shown to the subject and eye gaze tracking is done. For training and classification, we explored and compared three different approaches, namely:

1. Convolutional Neural Network (CNN)
2. Deep Neural Network (DNN)
3. Non-Neural Networks
    a. SVM
    b. Random Forest Classifier

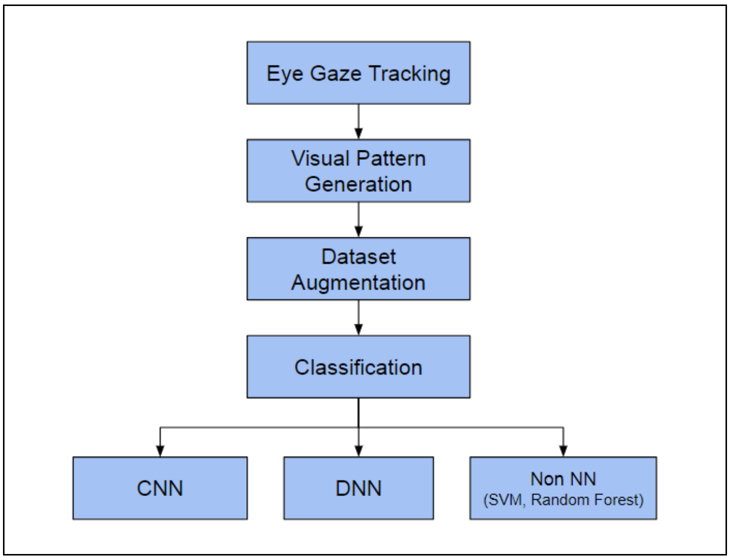


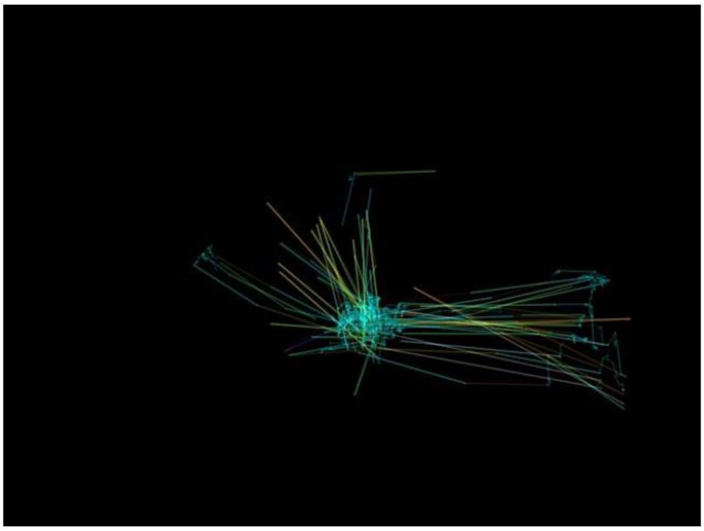
Fig. 5.1: Project work flow

5.1 Eye Gaze Tracking

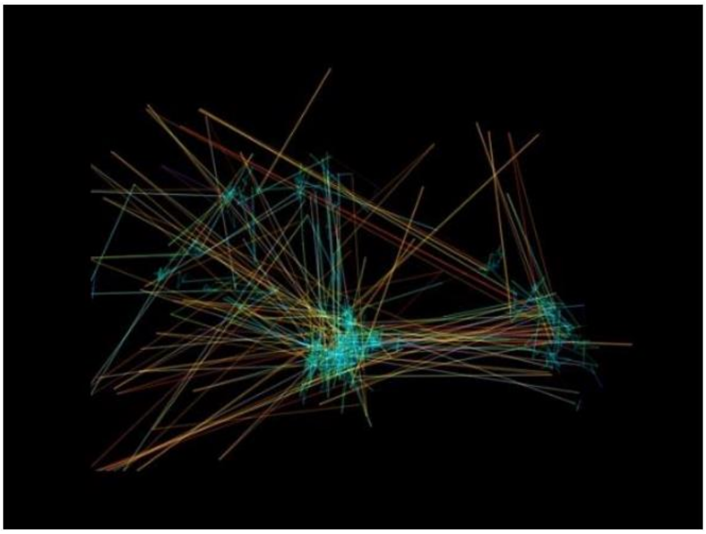
A one-minute video was shown to the subject and the gaze coordinates were recorded using iTracker[9]. The video includes a presenter who is speaking. The gaze coordinates were obtained as follows: First, face and eye detection was done on the input video stream using Opencv. The face, left eye and right eye images were cropped and a face mask was generated. The face mask is a binary mask which indicates the position of the face in the captured scene. The cropped face, left eye and right eye images were resized to 64*64 pixels and the face grid to 25*25 pixels. These are provided as input to the iTracker which generates the estimated gaze coordinates as output.

5.2 Visual Pattern Generation

A visual gaze pattern was generated from the obtained coordinates and saved as an image. A line is drawn for each transition from position [x(t), y(t)] to [x(t+1), y(t+1)], where t is a defined point of time during the experiment.The change in colour across the line represents the dynamics of eye movements. The RGB values were tuned based on the velocity, acceleration and jerk respectively at instant (t). This image is provided as input to the classification model. The implementation has been done using Matplotlib in Python.



(a)



(b)

Fig. 5.2 Visualization of eye gaze path (a) for a non-ASD participant (b) for participant diagnosed with ASD

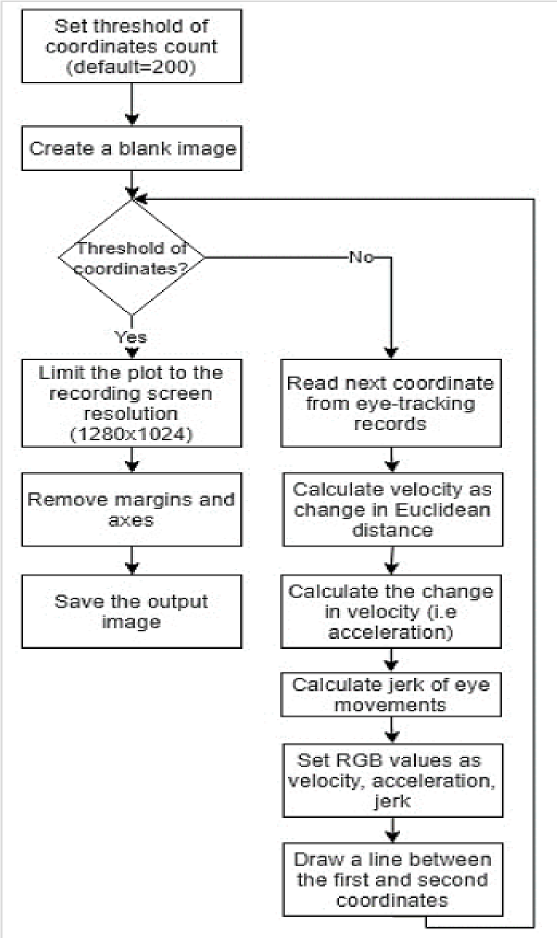


Fig. 5.3 Steps followed to visualize eye gaze tracking records

5.3 Image Augmentation

Image augmentation has been done to increase our dataset size to train Neural Network and Non-Neural Network models with limited or small amounts of data. We have used ImageDataGenerator class of Keras for image augmentation with random rotation range of 10 degrees, shearing range of 0.2 and zooming range of 0.2.

5.4Classification

We use image classification to classify eye gaze patterns as autistic or non-autistic. Classification can be done using either neural network approaches like CNN, DNN etc; or using Non-Neural Network approaches like SVM, Bayesian etc. In our methodology, we have used two Neural network based methods and two Non-Neural Network methods.

5.4.1 Convolutional Neural Network (CNN)

In deep learning, a convolutional neural network (CNN) is a class of deep, feed-forward artificial neural networks, most commonly applied to analyzing visual imagery.A CNN consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers and normalization layers.

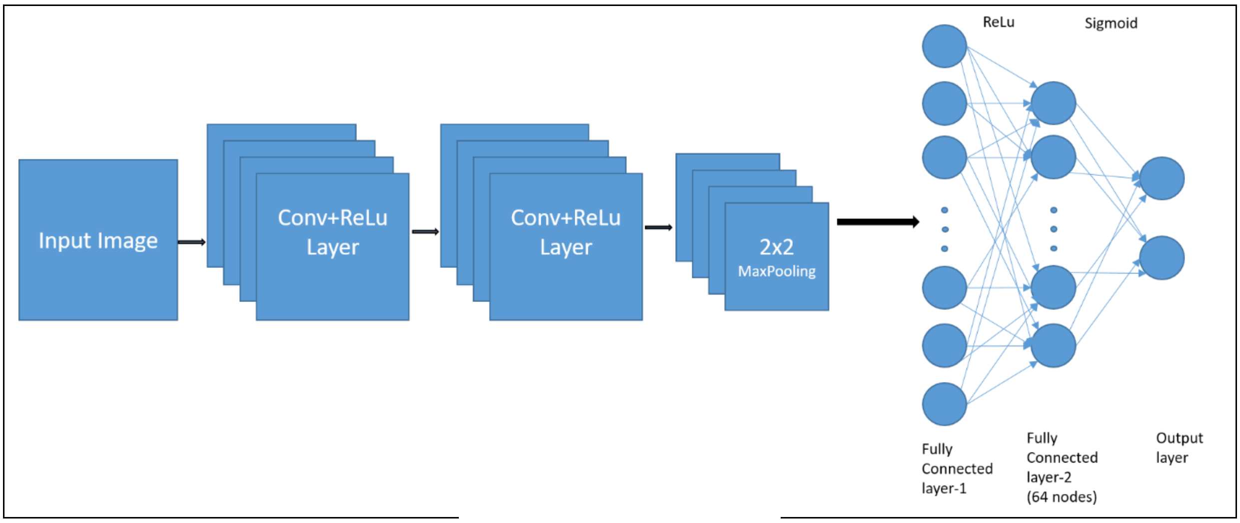


Fig. 5.4: Architecture of CNN

The basic architecture of our CNN has two convolutional layers followed by a MaxPooling layer, two fully connected layers and an output layer. The first two layers are convolutional layers that deal with input images that are seen as 2-dimensional matrices. Convolutional layers apply a convolution operation to the input, passing the result to the next layer. There are 32 nodes in each of the convolutional layers. The kernel size is 2, i.e., 2x2 filter matrix has been used as kernel. The activation function used is ReLu and shape of each input image is 200x200.

Next, Max Pooling layer has been used with pool size of 2x2. Max Pooling uses the maximum value from each of the 2x2 cluster of neurons at the prior layer and combines it into a single neuron in next layer.

Next, a flatten layer is used to connect convolutional layer and dense layers. The output is flattened to enter the Fully Connected Layers. First, we specify 64 nodes each activated by ReLu function. The next layer uses sigmoid function with binary classification to give the final output layer which consists of two nodes, one for each class.

Model has been compiled using Adam optimizer with learning rate of 0.005 and binary crossentropy as loss function. The model was trained for 1000 epochs with batch size 16. The ratio of training data and testing data was 70% and 30% respectively. Model has been implemented using Keras.

5.4.2 Deep Neural Network (DNN)

DNNs are typically feedforward networks in which data flows from the input layer to the output layer without looping back. The basic architecture of our DNN consists of four hidden layers and an output layer. Each set of 40000 input values act as input to first hidden layer. The first layer consists of 500 nodes and uses ReLu as activation function. The second hidden layer has 200 nodes and uses ReLu activation function. The third layer has 100 nodes and uses ReLu activation function. The fourth hidden layer has 50 nodes and uses ReLuactivation function. The next layer uses sigmoid activation function with binary classification to give final output layer which consists of two nodes, one for each class.

Model has been compiled on same parameters, i.e., using Adam optimizer with learning rate of 0.005 and binary crossentropy as loss function. The model was trained for 1000 epochs with batch size 16. The ratio of training data and testing data was 70% and 30% respectively.

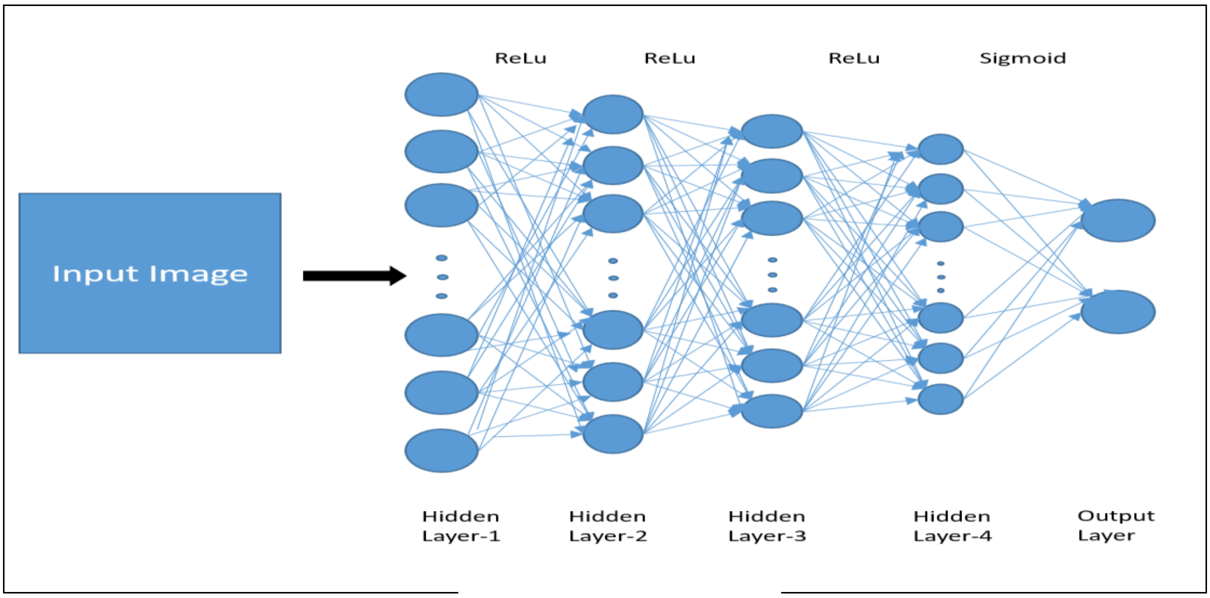


Fig. 5.5: Architecture of DNN

5.4.3 Non-Neural Networks

We have used two non-neural methods to train and classify eye gaze patterns. The first method used was Support Vector Machine (SVM) to classify eye gaze patterns as autistic or non-autistic. The second method used was Random Forest Classifier. Both of these are supervised learning algorithms.

6. RESULT

The neural network models were trained for 1000 epochs to determine the accuracy of the two approaches. The original dataset consisted of 219 eye gaze patterns of autistic children and 328 eye gaze patterns of non-autistic children. After image augmentation, the dataset size increased to 1043 eye gaze patterns for autistic and 1533 eye gaze patterns for non-autistic. The ratio of training data and testing data was 70% and 30% respectively, i.e., training was done on 1803 images, out of which 1081 images belonged to non-autistic class and 722 images belonged to autistic class and testing was done on 773 images. Figure 6.1 shows the overall accuracy of the two neural network methods, CNN and DNN, and two non-neural network methods, SVM and Random Forest Classifier.

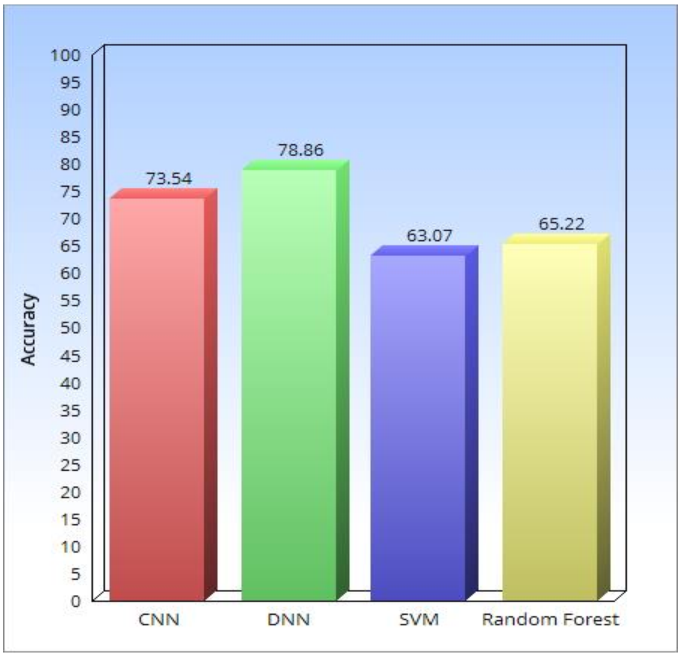


Fig. 6.1: Bar chart showing accuracy of the four methods

- Accuracy obtained using Convolutional Neural Network was 73.54% and Deep Neural Network was 78.86%.
- For non-neural network methods, the accuracy of SVM was 63.07% and Random Forest Classifier was 65.22%.
- The accuracy of neural network methods were better than that of non-neural network methods.
- DNN performed better than CNN.

Figure 6.2 shows the accuracy of autistic and non-autistic classes for CNN model and figure 6.3 shows the accuracy of autistic and non-autistic classes for DNN model.

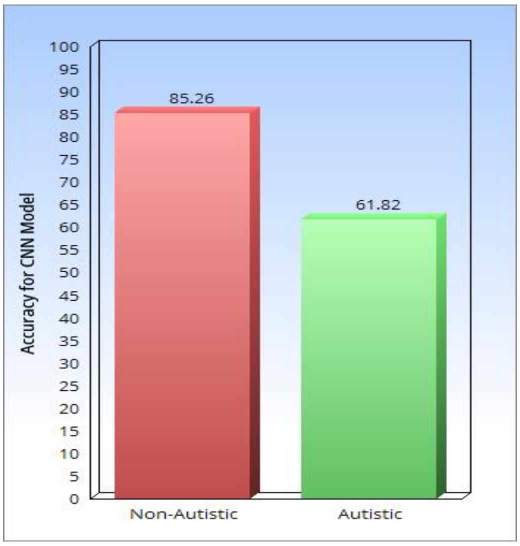


Fig. 6.2: Bar chart showing accuracy of the two classes for CNN Model

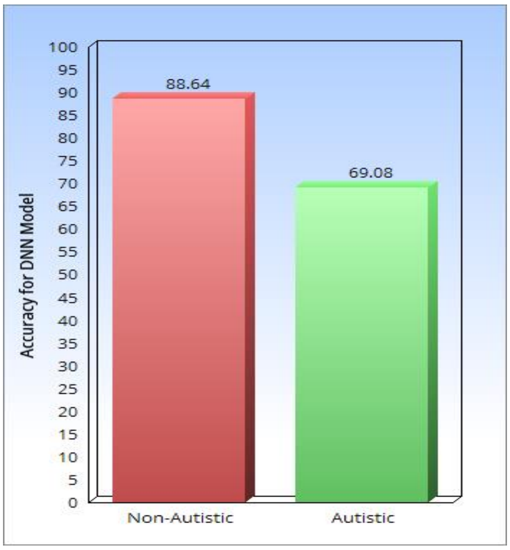


Fig. 6.3: Bar chart showing accuracy of the two classes for DNN model.

Table 6.1 and 6.2 show confusion matrices obtained using CNN and DNN.

	Non-Autistic	**Autistic**
Non-Autistic	85.26%	14.74%
Autistic	38.18%	61.82%

Table 6.1: Confusion matrix obtained using CNN

	Non-Autistic	**Autistic**
Non-Autistic	88.64%	11.36%
Autistic	30.92%	69.08%

Table 6.2: Confusion matrix obtained using DNN

The accuracy of autistic class is less than that of non-autistic class for both CNN and DNN because of lack of data from autistic subjects. The overall accuracy of DNN was better than CNN.

7.SOFTWARE & HARDWARE REQUIREMENTS

Implementation Language: Python 3

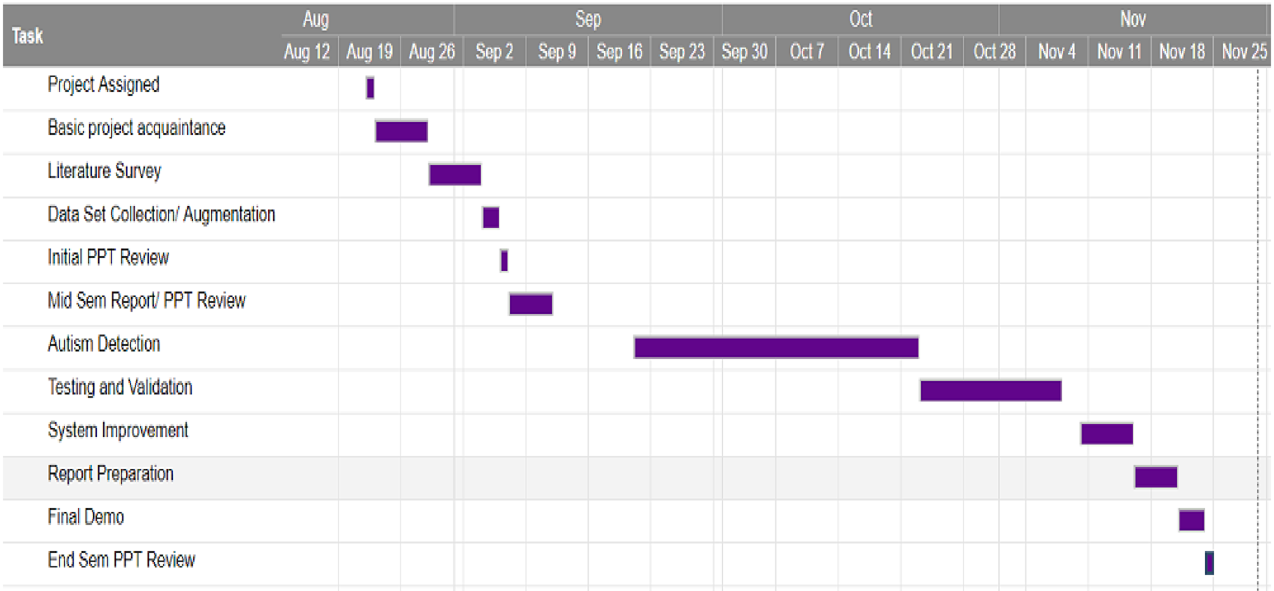
Libraries: Tensorflow, Keras, Matplotlib, OpenCV

Operating System Used: Ubuntu16.04LTS.

Hardware Requirements: 64-bit environment, intel Core i5 CPU @ 2.60 GHz with 8 GB of RAM

8.ACTIVITY TIME CHART AND GANTT CHART

TASK NAME	START DATE	END DATE
Project Assigned	22/08/2018	22/08/2018
Basic Project Acquaintance	23/08/2018	28/08/2018
Literature Survey	29/08/2018	03/09/2018
Data Collection/Augmentation	04/09/2018	05/09/2018
Initial PPT review	06/09/2018	06/09/2018
Mid sem Report/ PPT review	07/09/2018	11/09/2018
Autism detection	21/09/2018	22/10/2018
Testing and Validation	23/10/2018	07/11/2018
System Improvement	10/11/2018	15/11/2018
Report Preparation	16/11/2018	20/11/2018
Final Demo	01/12/2018	01/12/2018
PPT Review	01/12/2018	01/12/2018



9. CONCLUSION AND FUTURE SCOPE

We have trained three machine learning models for autism detection using eye gaze patterns. The methods used were CNN, DNN and two non-neural network methods, SVM and Random Forest Classifier. An accuracy of 73.54% was obtained using CNN, 78.86% using DNN and 63.07% and 65.22% for SVM and Random Forest Classifier respectively. This project also implements a functionality for transforming gaze coordinates into visual representation. Specialists (e.g. psychologists) can make use of such visual paths to assist the diagnosis.

The trained models can be integrated with an app or website to enable parents to get a pre-clinical diagnosis for autism in infants so that proper intervention can be provided for development of autistic children. Accuracy of the eye gaze tracking module needs to be improved. Accuracy of classification can be improved by training on a larger dataset.

10.REFERENCES

1. N. I. Vargas-Cuentas, D. Hidalgo, A. Roman-Gonzalez, M. Power, R. H. Gilman and M. Zimic, "Diagnosis of autism using an eye tracking system," 2016 IEEE Global Humanitarian Technology Conference (GHTC), Seattle, WA, 2016, pp. 624-627.

2. P. S. Rani, P. Subhashree and N. S. Devi, "Computer vision based gaze tracking for accident prevention," *2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave)*, Coimbatore, 2016, pp. 1-6.

3. R. YuliCristanti, R. Sigit, T. Harsono, D. C. Adelina, A. Nabilah and N. P. Anggraeni, "Eye gaze tracking to operate android-based communication helper application," 2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), Surabaya, 2017, pp. 89-94.

4. Boraston, Zillah & Blakemore, Sarah-Jayne. (2007). The application of eye-tracking technology in the study of autism. The Journal of physiology. 581. 893-8. 10.1113/jphysiol.2007.133587.

5. B. L. Nguyen, "Eye Gaze Tracking," *2009 IEEE-RIVF International Conference on Computing and Communication Technologies*, Da Nang, 2009, pp. 1-4.

6. Jie Zhu and Jie Yang, "Subpixel eye gaze tracking," Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, Washington, DC, USA, 2002, pp. 131-136.

7. Sewell, Weston &Komogortsev, Oleg. (2010). Real-time eye gaze tracking with an unmodified commodity webcam employing a neural network. 3739-3744. 10.1145/1753846.1754048.

8. Chong, Eunji&Chanda, Katha & Ye, Zhefan& Southerland, Audrey & Ruiz, Nataniel& Jones, Rebecca &Rozga, Agata&Rehg, James. (2017). Detecting Gaze Towards Eyes in Natural Social Interactions and Its Use in Child Assessment. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 1. 1-20. 10.1145/3131902.

9. K. Krafka *et al.*, "Eye Tracking for Everyone," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016, pp. 2176-2184.

Suggestions by Board Members