International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 05, 2020
ISSN: 1475-7192

Convolutional Neural Network for Prediction of Autism based on Eye-tracking Scanpaths

Zeyad A.T. Ahmed and Dr. Mukti E. Jadhav

Abstract--- *Autism Spectrum Disorder (ASD), difficulty in socialization, can be detected by observation of atypical visual attention of children. Eye tracking is one of the most important techniques used in providing information on visual behavior as a statistically-motivated step towards the accurate diagnosis of such disorder. The scanpath, sequences of fixations of the eyes on image, provides data related to the locations and durations of the gazes that can be used to develop visual patterns to analysis the visual behavior of children. The aim of this paper is to develop a deep learning model implementing a convolutional neural network (CNN) to classify children to autistic and typically-developing according to eye tracking scanpaths. The model was applied on 29 autistic and 30 normal children and achieved 98% testing accuracy.*

Keywords--- *Autism Spectrum Disorder (ASD), Typical Developing (TD), Eye Tracking, Scanpath, Convolutional Neural Network (CNN).*

I. INTRODUCTION

Autism spectrum disorder (ASD) is a neurological and developmental disorder that has a negative impact on an individual's learning, social interaction and communication. The diagnosis of ASD poses a great deal of difficulty as it needs a long term tracking and observation of the behavior of autistic children [1]. The signs of ASD appear in the early stage of children's life, but their diagnosis might be challenging especially in less than three-year-old children. Nowadays, the eye tracking technology helps to access the right diagnosis of Autism based on the visual patterns. The eye tracker devices, such as Tobii, SIM and Eyelike, are used to record and analyze eye movement. Fixation, Saccade, heat maps and scanpaths are examples of data collected for the analysis. This paper considers scanpaths of eye tracking in two groups of children, viz. 29 children with ASD and 30 typically-developing (TD) children, with the average age of about 8 years.

Eye tracking scanpaths provide information about how ASD and TD children visual attention respond to dynamic and static stimuli visual such as human activists and static images of different objects, which are shown to children during the recording session. SMI eye-tracker is used to capture the eye movement and record the fixation and saccade of the eye on the object. The data collected are used to build visual patterns of ASD as well as TD children. The visual patterns model should be able to identify whether a child has ASD or not based on scanpaths of eye tracking. This paper reports on the developed deep learning model based on a convolutional neural network (CNN), which is intended to help psychologists to get the right diagnosis of autistic children and then decide which treatment plan should be followed in order to improve the social communication of the autistic.

Zeyad A.T. Ahmed, PhD Research Scholar, Department of Computer Science, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India. E-mail: zeyad.ahmed2019@yahoo.com

Dr. Mukti E. Jadhav, Prof & Principal of Marathwada Institute of Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India. E-mail: muktijadhav@gmail.com

DOI: 10.37200/IJPR/V24I5/PR201970
Received: 23 Feb 2020 | Revised: 18 Mar 2020 | Accepted: 30 Mar 2020

International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 05, 2020
ISSN: 1475-7192

Section 2 of this paper covers some recent and relevant studies to the topic, for example deep neural networks and eye tracking scanpaths. Section 3 deals with the study methodology providing a detailed explanation of the dataset and CNN model for ASD detection along with the presentation of the result and analysis of the data. Section 4 discusses the study conclusion.

II. RELATED WORK

Convolutional neural networks become one of the most effective techniques for image classification especially with the potentiality of processing a large dataset. A few studies incorporate deep learning in autism detection based on eye tracking scanpaths and the most significant of them are going to be reported here.

Carette, et al. [2] developed visual patterns of eye tracking scanpaths model to detect autism by applying Artificial Neural Network (ANN) models into a dataset of 59 participants [3]. The result of the classificatory algorithm is 90% whereas the limitation of the model is the small number of participants. Elbattah, et al.[4] developred an unsupervised Machine Learning model to clustering autistic children based on eye tracking scanpaths with Deep Autoencoder. They used K-Means clustering algorithm using this dataset[3]. Goldberg, et al. [5] studied the scanpath representations of eye tracking and provided software for using eye tracking scanpath with web pages in which they record people visual attention and capture fixation form the beginning to the end of the session with the position in a sequence. The study describes how the scanpaths represented in axis (x, y). Rutherford, et al. [6] tested the scanpaths of two groups of people with ASD and TD using eye tracking. The Stimuli uses 20 images of female face with different emotions such as angry, surprised, happy, etc. The result reveals that the ASD people look at the eyes region less and are more interested in looking to the mouth.

Jiang, M., et al. [7] developed a method to classify groups of ASD and TD people based on Deep Neural Network (DNN) and support vector machine (SVM) using eye tracking data. The dataset consists of images eye tracking with heat maps and eye fixation. The classification of the feature is based on Gaze, and VGG-16 features. The accuracy level of this model is 92%. Wu, et al. [8] developed a deep learning model to predict autism using synthetic saccade and image fixation maps. Their dataset is 300 images of ASD and TD children recorded by Tobii T120. The images consist of scanpath data with a location and a duration. The first model is based on synthetic saccade using scanpaths with STAR-FC while the second is based on image fixation maps. The resultant accuracy validation of this dataset is 65.41%.

Tao, et al. [9] proposed CNN and long short-term memory (LSTM) networks to classify TD and ASD children, based on the scanpaths of eye gaze, a model using SP-ASDNet dataset, CNN-LSTM used to extract the features from the saliency maps and sequence of the fixation locations in the scanpath. The validation dataset accuracy of the model was 74.22%. Finally, Xie, et al.[10] develop a novel two-stream deep learning network to detect autism based on eye movement patterns of ASD and TD people. The dataset consists of 700 images. They transform VGG16 model to develop two VGGNets models which are able to classify ASD and TD. The accuracy result is 95%.

DOI: 10.37200/IJPR/V24I5/PR201970
Received: 23 Feb 2020 | Revised: 18 Mar 2020 | Accepted: 30 Mar 2020

International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 05, 2020
ISSN: 1475-7192

III. METHOD

In this study, the researchers aim to develop a deep learning model to detect autism based on a binary classifier making use of eye tracking scanpaths images of ASD and TD groups using CNNA. The proposed model Architecture is shown in fig.1.

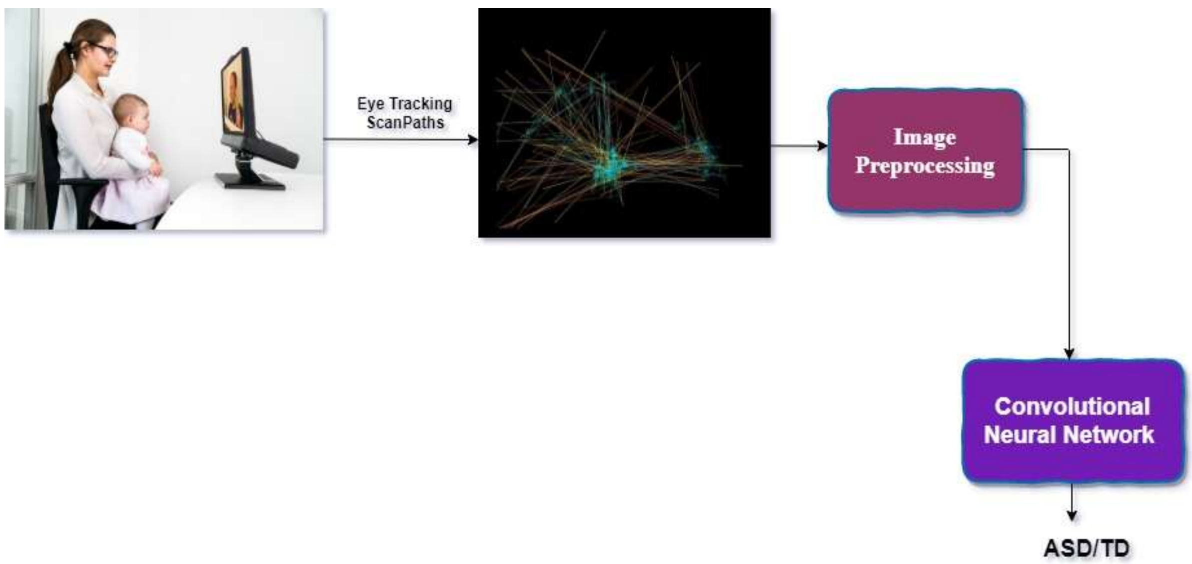


Fig. 1: The proposed Model Architecture

3.1 Participants

This study uses a publicly available dataset on the Figshare data repository [3] which was collected by Carette, et al. [2]. The study sample is 59 children participants who were tested by psychologists in special clinics in order to classify them into ASD and TD. As shown in fig.2, the classificatory distribution of the children maintains that 29 are ASD children (25 males and 4 females) while 30 are TD ones (13 males, 17 females). The average age of children is about 8 years old. The dataset contains 547 images; 219 images of ASD children and 328 TD children.

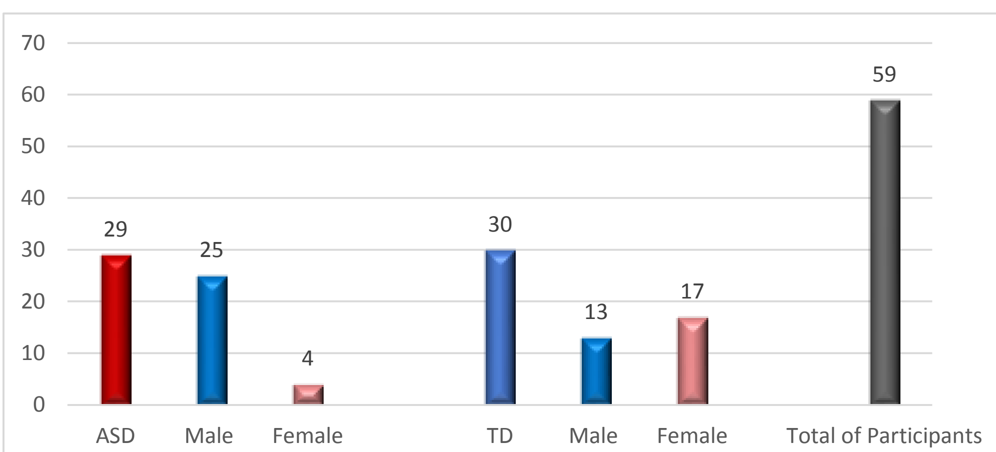


Fig.2: Number of Participants ASD and TD

DOI: 10.37200/IJPR/V24I5/PR201970
Received: 23 Feb 2020 | Revised: 18 Mar 2020 | Accepted: 30 Mar 2020

International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 05, 2020
ISSN: 1475-7192

3.2 Stimuli and Apparatus

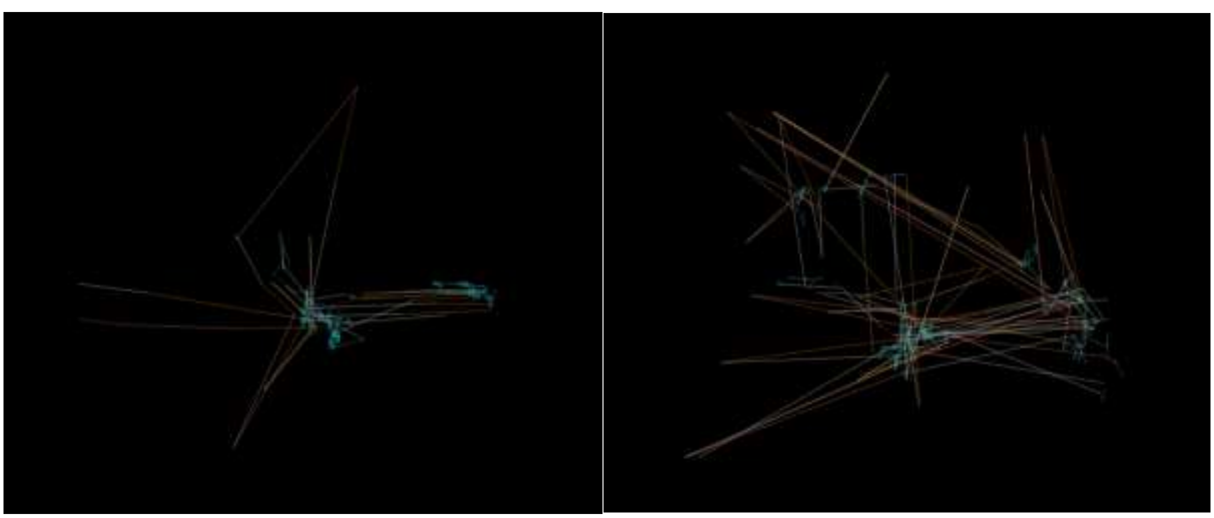
The data of eye tracking scanpaths are collected using SMI screen-based eye-tracker (Red-m 250Hz). The eye-tracker is situated at the screen bottom of a desktop PC or laptop. This experiment uses a 17-inch monitor of 1280x1024 resolution [2]. Two types of stimuli are utilized, namely, static stimuli and dynamic stimuli visual such as colorful cartoons, balloons as well as some videos that include a presenter who speaks French as the mother tongue of participants. SMI eye-tracker is used to capture the eye movement and record the fixation, saccade, blink on the object.

3.3 Experimental Procedure

In the eye tracking experiments, the experimenters collate the history of the children and some questions are posed to the child's parents [1]. The experiment room is dark and soundproof. The participant is seated in his/her parent's lap or sets by himself/herself in front of the screen. The distance between the screen and the participant is approximately 60 cm. Then she/he watches the dynamic visual stimuli. The stimuli consist of a set of videos to attract the attention of the child in order to record eye tracking scanpaths for developing visual patterns [1].

3.4 Visualization of Eye-tracking Scanpaths

Small images of scanpaths as connected saccades are used as scaled traces that allow the comparison and contrast of relative fixation densities and distributions of saccades. Eye tracking scanpaths provide information on how children perceive the static stimuli and dynamic stimuli visual. The scanpaths information are used to build visual patterns of ASD and TD. The visual patterns model will be able to detect whether the child has ASD or TD through scanpaths of eye tracking as shown in fig.3. Those images are from the dataset [3]. In this dataset, the participant scanpath of eye tracking can be recorded during the dynamics of eye motion and transformed into images. Those images are used to design visual patterns fixation for behavior analysis.



TD ASD

Fig. 3: Visualization of Eye-tracking Scanpaths

DOI: 10.37200/IJPR/V24I5/PR201970
Received: 23 Feb 2020 | Revised: 18 Mar 2020 | Accepted: 30 Mar 2020

3.5 Implementation

3.5.1 Image Augmentation

The dataset [3] consists of 547 images; 219 images of ASD children and 328 TD children images. The image augmentation is used to reduce the overfitting. Overfitting occurs when the samples in the dataset are small. The implementation of augmentation is to generate synthetic samples using a random set of image transformations such as rotation and shearing the images [2]. The image augmentation is applied to the dataset that consists of 547 images then each image becomes ten synthetic images.

3.5.2 Data pre-processing

In machine learning techniques, the first step is pre-processing of the collected data. It should be applied before training model. There are many pre-processing techniques such as read, remove the noise, images resize, segmentation and morphology (smoothing edges). In this dataset, [3] the image has to be resized and its color has to be changed to make the training model faster. The original size of images is 640*480*3 (640 high, 480 width, 3 is RGB). The image consists of 921,600 features [2]. When the number of image features have increased, the complexity of the training model and misleading may occur. To solve this problem, researchers resize the images by 150*150 and change the color of channels to a grayscale format so the input image to the CNN will be 150*150*1.

3.5.3 CNN Classification

CNN, a deep learning algorithm, takes the input as an image and assigns importance learnable weights and biases to various features in the image. CNN is able to detect the differentiating features in the images. In CNN, the pre-processing of the images is reduced compared to other classification algorithms. In traditional methods, filters are hand-engineered, with enough training while CNN can learn these filters.

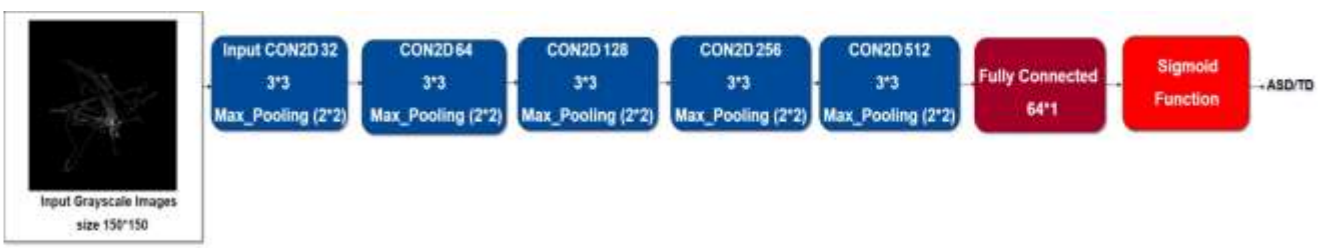


Fig.4: The CNN Network Architecture for Eye Tracking Scanpath of ASD and TD

The training process of the proposed ASD and TD classification model is performed on Lenovo laptop Core i5-3210M 2.50GH with NVIDIA GeForce 610M GPU. The platform is python version 3.7.3 with Tensorflow library version 2.1.0. The CNN architecture of the model contains 4 CONV2D Layers + 1 FC layer. The input layer is an image with size 150*150*1. The kernel and filter size are shown in fig.4. Each hidden layer is followed by Max_pooling filter with size 2*2. Each linear layer is followed by a relu activation function. The activation function of the output is sigmoid which is used to classify ASD and TD. The batch size is set to 32. Adam optimizer is used to train the networks with a learning rate of 0.001. The model is trained for 50 epochs and the dropout of the hidden layer is 0.20 to avoid overfitting.

DOI: 10.37200/IJPR/V24I5/PR201970
Received: 23 Feb 2020 | Revised: 18 Mar 2020 | Accepted: 30 Mar 2020

3.6 Experimental Result

This study develops different visual scanpaths patterns between ASD and TD children based on deep learning classification model. In this model, the data are divided into 70% for training and 30% for testing. The testing result of the model is 98% accuracy.

A confusion matrix is used to measure the prediction of the model. It is as a summary of prediction results on a classification of ASD and TD dataset. The table of confusion in fig.4 is used to describe the performance of the classification model. In this study, the researchers are building a classifier to predict whether the child has ASD or TD. The classifier is binary that contains two labels of prediction, namely, positive and negative. Positive means that a child has ASD while Negative indicates that the child is TD. Fig.5 shows the results of confusion matrix prediction of the model after it had been executed on 30% of testing data. It maintains the following details:

- **True positives (TP):** the model has correctly predicted 549 eye tracking scanpath images of the children who have ASD.
- **True Negatives (TN):** the model has correctly predicted, 487 eye tracking scanpath images of the children who are TD.
- **False Positives (FP):** the model falsely predicted 29 eye tracking scanpath images as ASD though they do not have ASD.
- **False Negatives (FN):** the model falsely predicted 10 eye tracking scanpath images as TD though they have ASD [11].

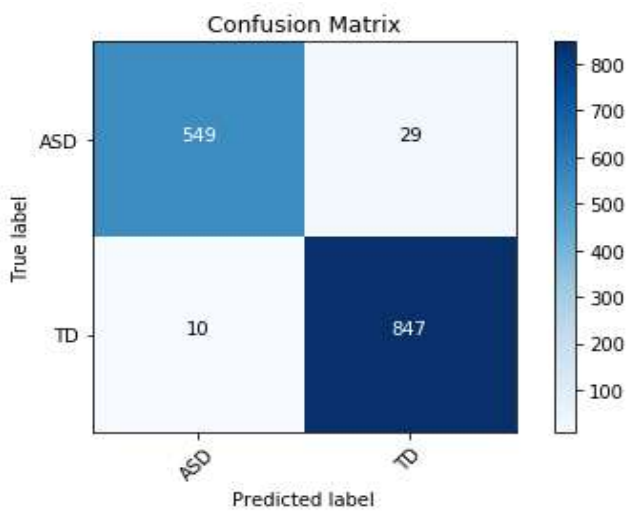


Fig.5: The Confusion matrix Table

IV. CONCLUSIONS

This paper presents a model of deep learning using CNN to classify the participants to ASD and TD based on eye tracking scanpaths. The model is applied to a publicly available dataset on the Figshare data repository [3]. The CNN architecture is a suitable algorithm to extract the features form eye tracking scanpaths images as a sequence of the fixation locations in the scanpath. In the testing of data, the accuracy achieved is 98%, which outperforms the best accuracy obtained on this dataset so far. The limitation of this study is the number of participants in the dataset [3], which is small. In the future, researchers are required to use more data to get a more robust model. The contribution of this work aids psychologists to diagnose children with ASD in the early stages of their life efficiently and easily. It is so because the diagnosis process of such psychological disorder is difficult and requires more time and effort especially to observe the behavior of the child. Therefore, more efforts have to be made to develop technological devices and models to do the job.

REFERENCES

[1] Zeyad Abdulhameed Taha Ahmed, Mukti E. Jadhav , A Review of Early Detection of Autism Based on Eye-Tracking and Sensing Technology, *5th International Conference on Inventive Computation Technologies (ICICT-2020), INDIA, IEEE.* "Under Publication"

[2] Carette, R., Elbattah, M., Dequen, G., Guérin, J., Cilia, F., & Bosche, J. (2019). Learning to predict autism spectrum disorder based on the visual patterns of eye-tracking scanpaths. *In Proceedings of the 12th International Conference on Health Informatics.*

[3] https://figshare.com/articles/Visualization_of_EyeTracking_Scanpaths_in_Autism_Spectrum_Disorder_Image_Dataset/7073087/1

[4] Elbattah, M., Carette, R., Dequen, G., Guérin, J. L., & Cilia, F. (2019, July). Learning Clusters in Autism Spectrum Disorder: Image-Based Clustering of Eye-Tracking Scanpaths with Deep Autoencoder. *In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1417-1420). IEEE.

[5] Goldberg, J. H., & Helfman, J. I. (2010, March). Visual scanpath representation. *In Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (pp. 203-210).

[6] Rutherford, M. D., & Towns, A. M. (2008). Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. *Journal of autism and developmental disorders*, 38(7), 1371-1381.

[7] Jiang, M., & Zhao, Q. (2017). Learning visual attention to identify people with autism spectrum disorder. *In Proceedings of the IEEE International Conference on Computer Vision* (pp. 3267-3276).

[8] Wu, C., Liaqat, S., Cheung, S. C., Chuah, C. N., & Ozonoff, S. (2019, July). Predicting Autism Diagnosis using Image with Fixations and Synthetic Saccade Patterns. *In 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)* (pp. 647-650). IEEE.

[9] Tao, Y., & Shyu, M. L. (2019, July). SP-ASDNet: CNN-LSTM Based ASD Classification Model using Observer ScanPaths. *In 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)* (pp. 641-646). IEEE

[10] Xie, J., Wang, L., Webster, P., Yao, Y., Sun, J., Wang, S., & Zhou, H. (2019). A Two-stream End-to-End Deep Learning Network for Recognizing Atypical Visual Attention in Autism Spectrum Disorder. arXiv preprint arXiv:1911.11393.

[11] Mohamed Elgendy, (2019) Deep Learning for Vision Systems available online on https://www.manning.com.