Classification of Autism Spectrum Disorder and Typically Developed Children for Eye Gaze Image Dataset using Convolutional Neural Network

Ms. Praveena K N, Dr. Mahalakshmi R
Department of Computer Science and Engineering
Presidency University, Itkalpur, Rajanukunte, Bengaluru, India

***Abstract*—Autism is a neurobehavioral problem that hinders to interact with others. Autistic Spectrum Disorder (ASD) is a psychological disorder that hampers procurement of etymological, communication, cognitive, social skills and Stereotypical motor behaviors and capabilities. Recent research revealing that Autism Spectrum Disorder can be diagnosed using gaze structures which has opened up a new field where visual focus modelling could be highly used. Diagnosis of ASD becomes a difficult task due to wide range of symptoms and severity of ASD. Deep neural networks have been widely employed and have shown to perform well in a variety of visual data processing applications. In this paper, typical developed (TD) or ASD is classified using Convolution neural Networks (CNN) for the fixation maps of the corresponding observer's gaze at a given image. The objective of this paper is to observe whether eye-tracking data of fixation map could classify children with ASD and typical development (TD). We further investigated whether features on visual fixation would attain better classification performance. The proposed CNN model achieves 75.23% accuracy for validation.**

Keywords*—Autism spectrum disorder; classification; fixation maps; eye expression; visual focus; gaze pattern; CNN

I. INTRODUCTION

Currently, the world is facing lot of difficulties in medical field to diagnose the diseases in early stage. Doctors and specialists did not have the benefit of technology to identify a disease in prior in order to take precautionary measures to predict the diseases in early stage [26]. Once the condition is in risky level and becomes too late for treatments, it couldn't help to diagnose the disease. Therefore, by using data science algorithms, the data can be analysed efficiently to gain meaningful knowledge about the status of a person's health. This increase in figuring technologies for the presentation of deep learning techniques in numerous grounds of learning. Here in this research I have taken ASD Eye gaze patterns image dataset for classifying autistic and typically developed children.

Humans have an amazing tendency to focus on certain aspects of an image rather than viewing the whole scene in its entirety. Simulating the Visual Focus Modelling (VFM) selective focus function, also known as visual focus prediction or visual saliency detection. In the fields of computer vision and neuroscience, this is a well-known research subject. Visual attention helps observers to recognise the essential regions of a scenario, which is a critical aspect in many applications. Various This research is useful in the development of visual attention models such as bottom-up and top-bottom saliency models, saccadic models, and models for identifying items of interest.

Recently, several studies have shown that gaze characteristics can be used to recognize emotional conditions, perceptive processes, and neuropathology particularly in persons with Autism Spectrum Disorder (ASD). This has opened up a new field in which visual focus modelling can be beneficial in a variety of ways, including assisting in the early identification of ASD and building Computer-Human Interfaces (CHIs) that are suited for people with ASD.

It has been found that youngsters with ASD can have uncommon examples in gaze perception, which is affected by the kid's essential visual handling being unfocussed.

In this paper, the fixation maps which represents the locations of the gaze pattern responding to the given stimuli which is useful to decide whether the observer has ASD or not [3]. Therefore, a Grand Challenge named "Saliency4ASD: Visual centre demonstrating for Autism Spectrum Disorder" will be held at IEEE ICME'19 Grand challenge. It is one of the Dataset Repository and the key goal of this initiative is to coordinate and guide the visual focus modelling community's activities toward a social challenge in healthcare. The dataset containing fixation maps of children with ASD and TD for 300 images were shared as part of the Saliency4ASD grand challenge in 2019.

This paper builds a classification model using deep learning technique called convolutional neural networks (CNNs) model for determining if an observer is typical developed (TD) or has ASD based on the fixation maps of the associated observer's gaze at the captured image. The objective of this paper is to observe whether eye-tracking data of fixation map could classify children with ASD and typical development (TD). We further investigated whether features on visual fixation would attain better classification performance. The rest of the paper is organized as follows: Section 1 deals with introduction, Second Section gives a detailed literature review on classification of ASD, in particularly, the works connected to the visual attention study from a cognitive perception. Section 3 briefing about Autism Spectrum Disorder. Autism is a neurobehavioral problem that hinders to interact with others. Autistic Spectrum Disorder

(ASD) is a psychological disorder that hampers procurement of etymological, communication, cognitive, social skills and Stereotypical motor behaviours and capabilities. Section 4 describes the thorough overview of the Grand Challenge "Saliency4ASD operation. Sections 5 and 6 discuss the datasets, its types, methods & metrics used in the proposed work. Section 7 and 8 represent the proposed model for classifying ASD/TD and Results respectively.

II. Related Work

Almost all the people with ASD are experiencing more challenges with social communication and association, and finding difficulty in showing confined examples of behaviour, interests, or exercises in any given task [9]. The drawn out related issues, may incorporate troubles in performing every day works, making and keeping connections, and looking after a basic work. ASD isn't just perhaps the most confounded mental issues, yet additionally it is quite possibly the most effortlessly acquired. A more experienced parent, an ASD family history, and some inherited disorders are all risk factors for ASD. Because of the mind boggling quality to-quality and quality to-climate cooperation systems, some basic investigations demonstrated that the genetic and atomic premise of the people with ASD include in excess of 100 qualities [23]. There is a reasonable differential between those with ASD and those with Typical Development (TD), particularly as far as neighbourhood wisdom instead of worldwide perception.

People with ASD pay more attention to neighbourhood details with a more grounded tactile capacity, according to certain related works as well as improved perceptual discrimination along with an extraordinary tactile reactivity, such as eruption or under reaction to the climate [1].

At the moment, the most common methods for detecting ASD are conducting interviews with guardians or family members of ASD affected children, as well as observation and examination of their behaviour. Nonetheless, these finding methods are not just powerless against the inclination of the abstract clarification, yet in addition are time-devouring and costly. Furthermore, due to numerous limits and blockages, more than a third of ASD testing does not follow the precise and accurate diagnosis pathway [24]. A few recent research have recommended for the use of the aggregate technique to overcome the difficulties of genotyping investigations on complex mental diseases such as schizophrenia. The Broad Autism Phenotype (BAP) was proposed by few other studies few surveys have discovered that the parents of children with ASD are gifted individuals. At the same time, these parents were experiencing ASD symptoms unknowingly.

Dataset serves as a substantial aggregate to aid in the diagnosis of ASD development. Eye tracking data comprises vast data of visual consideration, neurological control, and individual psychology, Furthermore, eye development data has been successfully employed to assist with several other mental turmoil conclusion tasks, such as mental state acknowledgment and neuropathology determination [24]. In contrast to old-style quantization table-based determination procedures, eye development information-based strategies are unbiased, dependable, efficient, and logical due to the programmed assortment and quantitative processing. As a result, image statistics on eye development can greatly aid in ASD diagnosis. Few studies have discovered that people with ASD have different personality’s eye movement behaviours than people with normal. In terms of preference predispositions, the study and evaluation of eye development records may afford a technique to identify between people with ASD and those with TD [25]. ASD kid’s, for example, give a smaller amount of attention towards the frontal area entertainers and pay more attention to the ground districts when viewing a movie. Young people by ASD give a smaller amount of attention towards the eyes of human faces than those with TD. While watching recordings, Jones et al. [13] used technology of eye tracking to investigate the dispersal of optical obsessive information in people with TD and those with ASD. According to examinations, people with ASD give a smaller amount of attention to the entertainer’s eyes and give extra responsiveness to their mouths [18]. Additionally, there have been current research improving this problem that use AI propels and a deep learning method has been taken to anticipate where children with ASD will glance in various settings [5], with a focus on human looks. Fang et al. [8] also recommended using gaze following associated visual enhancements for autism analysis.

Advances in eye-development information studies show that people with ASD have different eye-development behaviour than those with TD.

III. Autism Spectrum Disorder

Mental imbalance range issue is an assortment of multipart issues of the mind advancements remember trouble for social connections, dreary practices and interchanges. These issues root at youthful age of 2-3 years and thusly, indications of ASD might be acknowledged at a youthful age [6] Autism Spectrum Disorders are multiple times regular in young men than in young ladies [5].

A. *The Causes of Autism*

The causes of autism are as follows:

1) It can be happened in children of any race, civilization, or communal family.

2) Chemical imbalance may cause because of specific blends of qualities which may build a youngster's danger in chemical imbalance.

3) Chemical imbalance is more likely in a child with a more experienced parent.

4) On the off chance that a pregnant lady is presented to specific drugs or engineered synthetic substances, like liquor or hostile to seizure prescriptions, her kid is probably going to be therapeutically contemplative.

5) Other danger factors incorporate maternal metabolic issues like as diabetes and obesity. Untreated phenylketonuria (otherwise called PKU, a metabolic contamination affected by the shortfall of a synthetic and rubella) has additionally been connected to psychological instability (German measles).

B. Autism Screening and Diagnosis

It is hard to get a definite conclusion of chemical imbalance. The specialist will accentuation on exercises and development of youngsters. Diagnosis of Autism can takes place in two methods:

- A progressive screening will verbalize the specialist whether the youngster is on target with rudimentary capacities, for example, schooling, talking, exercises, and moving. Pros suggest that kids be screened for these formative deferrals all through their precise registration at 9 months, year and a half, and 24 or 30 months old enough. Kids are constantly checked decisively for mental imbalance at their 18-month and two year registration.
- In the event that the youth exhibits signs of an issue on these screenings, they will require an additional an unmistakable appraisal. This might incorporate hearing and vision appraisals or inherited tests. The expert might need to bring somebody who has useful involvement with mental awkwardness issues, like a developmental pediatrician or an adolescent clinician. Picked investigators can other than give a test called the Autism Diagnostic Observation Schedule (ADOS) [21].

C. Autism Treatment

Here is certainly not remedy aimed at chemical imbalance. However, initial determination be able to gain an extraordinary change in the ground for a kid with mental imbalance. There are two main treatments of autism. They are:

- Behavioural and open treatment are two sorts of treatment. One of these medications is Applied Behaviour Analysis (ABA), which advances positive conduct while restraining negative conduct.
- Behavioural and open treatment are two sorts of treatment. One of these medications is Applied Behaviour Analysis (ABA), which advances positive conduct while restraining negative conduct.
- Sensory joining treatment can assist the individuals who with having issues with being reached or with seeing or hearing things.
- Speech treatment creates social limits.

IV. Summary of the Dataset

During ICME'19, Dell, Invensun, and "Quest Industries Creative" co-supported the Grand Challenge "Saliency4ASD [1]: Visual thought introducing for Autism Spectrum Disorder," which was directed by Shanghai Jao Tong University and the University of Nantes (a Research, Education and Innovation gathering of the Region Pays de la Loire, France).

Major points of these events existed as follows:

- Change visual consideration nearby district to a clinical consideration social issue, just as to help field experts in building a first benchmark of models that might be used in the analysis of ASD and the progression of gainful apparatuses for people with ASD. This Grand Challenge's underlying technique focused exclusively on cutting edge youngsters with ASD and kids with TD, disregarding the chance of comorbidities (i.e., when no less than two issues co-happen in a comparative subject, like Attention Deficit Hyperactivity Disorder) and the different degrees of ASD inside the synthetic awkwardness range [9]. By taking a dataset of pictures, cover a wide-ranging scope of substance and together with total eye-following data got through visual thought tests with kids with ASD and with TD [1]. This dataset will take into account the preparation and tuning of visual thought calculations, just as a reasonable correlation of their shows.
- Standard models used for saliency and grouping, given that devices, and determining appropriate measurements are all available in this grand challenge [1].
- Identify probable concerns from the datasets that can aid in the identification of individual look examples using autism, for instance the impact of the content (like photographs with humans, without humans, some human faces, and so on.
- Given that a pattern of exhibiting methodologies on behalf of ASD [1], which give intuitions of the essential highlights powerful optical consideration of persons with ASD [8], along with thoughts for probable demonstrating systems (like machine learning, training techniques, network models [8], and so forth).
- Considering this, the proposed experiments meant for the observers mounted as:
  - Track 1: Assumed an image, calculate the saliency maps that [1] correspond to people with ASD's gaze behaviour.
  - Track 2: Given an image and one observer's fixation sequence, classify whether he or she has ASD or TD.

V. Description of the Dataset and the Processed Eye Gaze Data

The dataset made for the Saliency4ASD Grand Challenge, which contains pictures and eye tracking information of youngsters with ASD and TD, is displayed in this section.

A. Experiment with Subjectivity

To create a dataset for analysing observation forms in kids with ASD, a one-on-one experiment was conducted with kids with ASD and TD [20], in which they were educated to openly observe pictures however their eye movements were captured using an eye-tracker [1].

We procured 300 pictures from, which is an enormous public data base that contains pictures with different scenes [4], to concentrate on the characteristics and changes in eye movements among kids with ASD and ordinary children under various visual improvements. The test boosts incorporate 40 pictures of different animals, 88 pictures of structures or articles, 20 pictures of typical scenes, 36 pictures of numerous people in a solitary picture, 41 pictures of different individuals and items in a solitary picture, 32 pictures of a solitary person in one picture, and 43 pictures of a single individual and things in a single picture [4].

Images were displayed and eye movements were recorded using a Tobii T120 Eye Tracker [17]. This eye tracker has a 17-inch display with a 1280*1024 (width*height) resolution [4]. The eye tracker's sampling rate is defined as 120 Hz, and the tracking range is 50 to 80 cm [4]. Subjects are positioned at a viewing distance of 65 cm from the eye tracker [4] in our studies.

B. Subjects

Twenty exceptionally effective kids with ASD who satisfied DSM-V [4] analytical principles for autism [10] were included in the dataset of eye movements for kids with Autism Spectrum Disorder (ASD) [19][13][14]. Amongst 20 members with ASD, just 14 subjects might finish the adjustment stage then get active eye development information [15]. The average age for the participants were 8 years. Along with this ASD, 14 healthy kids were recruited as controls, whose average age is 8 years [4]. In the dataset, gender, race and education were matched with two groups to guarantee the speculation of the data set. In the given dataset, parents gave written consent.

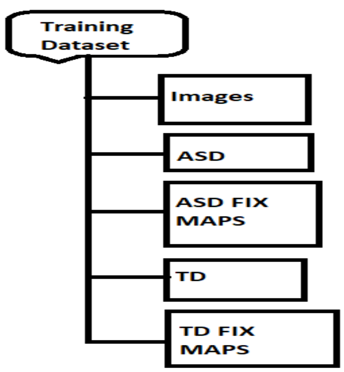


Fig. 1. Training Dataset Structure.

The Grand Challenge participants were given a training dataset in response to their request [1]. It indicates that this dataset is freely accessible for research scholars [1].

This dataset was provided to members after they requested it in order to construct and tune models. It contains 300 images as well as eye tracking data from 14 children with ASD and 14 children with Typical Development (TD) [16]. As may be seen in [2] this dataset is open to the research community. Fig. 1 represents the structure of Training dataset which is described as follows:

- Images: Images are in the form of “png” format.
- ASD: Text documents containing children with ASD's fixation sequences. Each document is a single image with fixations from all individuals [1]. It shows the index of fixation, as well as the A and B coordinates [1] and fixation time [1]. The fixation index [1] (Idx) goes from 0 to 1 [1].
- ASD Fixmaps: This record contains the saliency maps (in "png" documents) of each picture got from fixations of youngsters with ASD [1], following separating with a 1-degree Gaussian channel [1].
- TD: Manuscript documents that contain the fixation structures of children with TD.
- TD Fixmaps: This record has the saliency maps [1] (in "png" documents) of every picture obtained from TD fixations [1].

VI. Methods and Metrics

The methodologies and metrics used to assess the models performance are discussed in this section. The following metrics could be used for comparing saliency maps to predict gaze pattern of ASD children.

Accuracy: Accuracy is defined as the number of correct predictions divided by the total number of input samples [11].

Accuracy = (Correct number of guesses) / (Number of total predictions made)

F1 score: It is used to assess the test's precision. The Harmonic Mean of recall and precision is the F1 Score. F1 Score has a range of [0, 1] [5]. It lets you know how explicit and robust your classifier is (the number of occurrences it accurately arranges) (it doesn't miss a critical number of examples).

Mathematically, it is written as:

F1=2* 1/((1/precision)+(1/recall))[17]

The F1 score determines the stability of precision and recall [5].

Precision: The number of correct positive results divided by the number of positive results predicted by the classifier is the precision [5].

Mathematically, it is written as:

Precision=True positive/(True positive+False positive)[16][11][22]

Recall: It is calculated by dividing the number of correct positive results by the total number of relevant samples [5] (all samples that should have been classified as positive) [5].

Recall=True positive /(True positive+FalseNegative) [8]

VII. Proposed Model for Classifying ASD and TD

A. Methodology

Fig. 2 depicts the flow chart of research methodology and it tells about the data collection, data pre-processing & augmentation and classification. The approach describes the method that will be used to carry out the experiment. It entails data preparation and enhancement, classification, prediction using CNN model and performance evaluation.

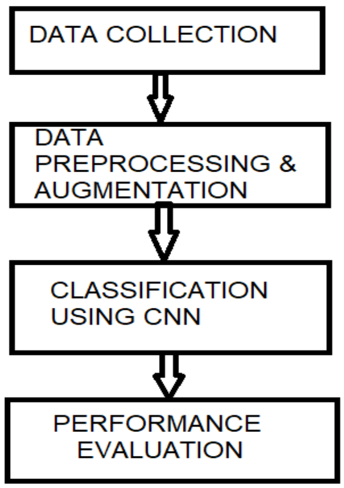
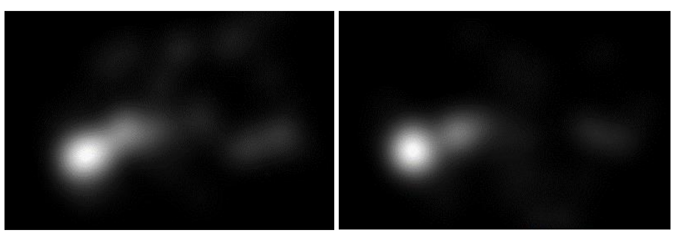


Fig. 2. Flow Chart of Research Methodology.

Fig. 3(a) and (b) denotes sample image of ASD and TD fixation map respectively. This research involved the automatic detection and classification of ASD and TD for ASD and TD fixation maps [2]. First the input pictures are pre-processed and features are extracted automatically using Convolution Neural Networks (CNN). The CNN architecture is used to classify the ASD and TD. Sample example of ASD and TD fixation maps are as follows:



(a) (b)

Fig. 3. Sample Image of (a) ASD Fixation Map (b) TD Fixation Map.

ASD Fixmaps: This record contains the saliency maps (in "png" documents) of each picture got from fixations of youngsters with ASD [1], following separating with a 1-degree Gaussian channel [1].

TD Fixmaps: This record has the saliency maps [1] (in "png" documents) of every picture obtained from TD fixations [1].

For classification challenges, deep learning technology is applied. When compared to traditional architecture, it processes a huge number of hidden layers. It learns features by consuming a huge quantity of labelled data without using any feature extraction methods [2]. CNN is a core deep learning algorithm that gives the best results on a variety of databases, including Cifar-10, MS-coco Mnist, and others. CNN is utilised in this study to classify ASD and TD using fixation maps.

B. *Convolutional Neural Network (CNN)*

CNN is a feed forward neural network [17][6] that is mostly used to analyse computer vision problems.

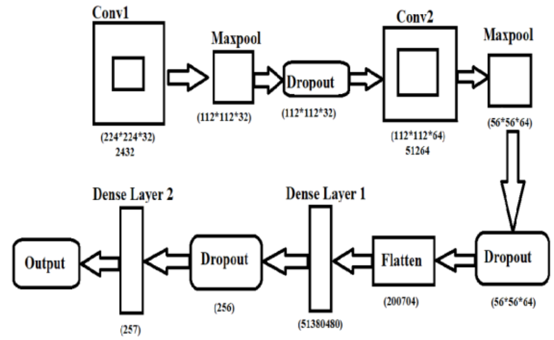


Fig. 4. Flowchart for CNN Implementation.

Fig. 4 depicts the CNN implementation flowchart. It consists of Convolution layer followed by Maxpool layer, dropout layer, flatten layer which is followed by dense layers. The architecture is represented by a multi-layered perceptron with shift invariant properties. Multiple layers are interconnected and constructed based on the human visual brain. The input layer, hidden layer, and output layer are the three main layers that make up the CNN architecture [3]. The input layer [3] represents about the image that will be entered, it pre-processes the image to make it uniform in height, width, and channel count. Convolution is depicted in the second major layer. With stride and padding, an array of uniform filters is convoluted to the input images in this layer. The equation defines the convolution layer for this work.

$$D_{i,i,j} = \sum_{k=0}^{k-1} \sum_{p=0}^{H-1} \sum_{q=0}^{H-1} I_{i+p,j+q,k}^{l-1} h_{pqkm} + b_{ijm} \quad (1)$$

Where Di,i,j is the convolution's output, and Fpqkm represents weight applied to the convolution, and bijm represents bias applied to the convolution [3]. I=1...mandj=1...n represents row and column indexes of input image [3], Fpqkm represents weight applied to the convolution [3], and bijm represents bias applied to the convolution [3]. Equation (1) is used to calculate striding and padding.

([W-F+2P]/S) +1 (2) [11]

Where:

W is the number of input volume size [11],

F is the size of filters to be [11] convoluted with input,

P is the padding and.

S is the stride.

In deep learning approaches, the Rectified linear unit (ReLU) [7] is a widely used activation function [15] [7]. It is complex to enhance weight through the gradient descent if the input has narrow derivative, ReLU allows to reduce the vanishing gradient problem. ReLU activation function is derived by the following equation (3).

(di,k) =max(0,Di,k) (3)

(x) = αx for x<0

x for x≥0

Where (x) denotes ReLU parametric function and $(i,)$ denotes convolutional output from the filter and diseased image. Pooling layer serves to reduce the spatial resolution of convolutional outcome on each dimension. Pooling has designed using maximum, minimum or mean values of the kernel in an image. Max pooling is widely used for convolutional neural network.

Let K x K block is taken from the NXN image. Maximum value is selected from KXK block and placed it in a pool. Commonly 2X2 window is preferred for pooling layer.

125	179	166	176
174	278	263	225
147	242	233	225
129	187	175	209

(a)

278	**263**
242	**240**

(b)

Fig. 5. Pooling Layer (a) 4*4 Convoluted Output (b) 2*2 Max Pooled Output.

The values from the convolution layer are shown in Fig. 5(a), and the max polled values from the specified convolution layer are shown in Fig. 5(b).

After collecting features from the convolutional layer, the outputs are processed through a pooling layer before being flattened into a single-dimensional vector. A 2X2X64 convoluted 2D vector is converted to a 256 flat vector using the flatten filter. The flattened layer's output is connected to the activation function by the dense layer. Each input node is connected to its corresponding output node. The data in the dense layer is normalised using the ReLU activation function, where wk is the weight function present in the kth node of [3] the flattened layer. In this study, CNN is used to do binary classification of ASD and TD. The trained model for ASD classification was created using the Keras model's prediction class, which accepts trained data as Numpy Arrays as input. Binary classes are employed in this study, hence binary inputs have been converted to arrays as dictionaries. Class weights are assigned to [3] each sample's keys in the training [3] phase. For each label, the sigmoid activation function is utilised to categorise the probability distribution and which is shown in Fig. 6.

Sigmoid function is given by:

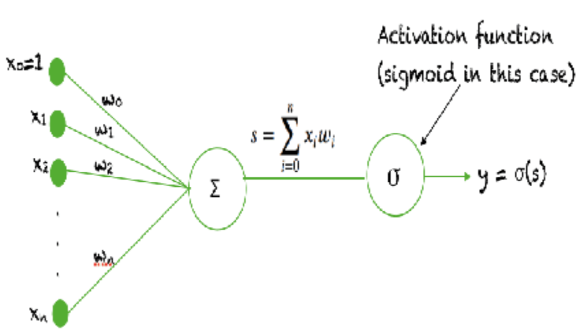


Fig. 6. Sigmoid Activation Function.

At the point when the activation function for a neuron is a sigmoid function, it is an assurance that the yield of this unit will consistently be somewhere in the range of 0 and 1. Likewise, as the sigmoid is a non-linear function, the yield of this unit would be a non-direct function of the weighted amount of data sources. Such a neuron that utilizes a sigmoid function as an activation work is named as a sigmoid unit.

C. *Proposed Model for Classifying ASD and TD*

Fig. 7 represents the CNN implementation for the proposed model. By using ASD/TD fixation maps as an input image, the proposed methodology explains how to classify ASD and TD. Deep neural networks are used to extract hidden information from fixation maps and make decisions in this case [3]. Our proposed classification model, which is based on the CNN architecture [2] and uses the fixation maps, is shown in Fig. 6. Fig. 6 shows our proposed categorization model based on eye tracking fixation maps. Fixation maps are incorporated into a proposed model to classify whether the person is having autism or not [2] [3]. The eye movement dataset [2] for ASD children [2] provided by the Saliency4ASD grand challenge organiser is used to train the ASD classification networks [2] [3].

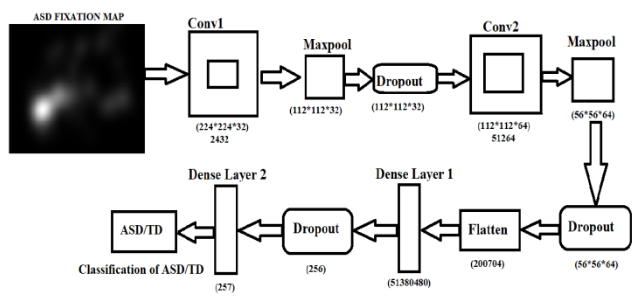


Fig. 7. Proposed Model for CNN Implementation.

VIII. Experiment

A. *Dataset Usage*

The training dataset comprises 300 images, where every image is observed by 14 TD youngsters and 14 ASD kids [3]. Meanwhile a few of the youngsters probably won't take a gander at the areas inside the pictures, a portion of the pictures have under 14 fixation maps from the ASD or TD gatherings.

The first 70% of images (images 1-210) and their associated fixation maps are used for training the proposed model [2] in the Saliency4ASD grand challenge, while images 211-300 and their associated fixation maps are utilised as the validation dataset to [2] choose the model for testing [2].

B. *Experimental Setup*

The proposed ASD classification model's training process is carried out in Google colab. The Adam optimizer [2] is utilized to train any deep learning model [2], with a learning rate [2] of 0.001 and a batch size of 20. The model is prepared for [2] 60 epochs, then the model that performs best on the validation dataset [2] is picked for testing. In the time of the evaluation, the model's weights are stable [2], and dropout is [2] utilized.

IX. RESULTS

The major goal of this research is to use Fixation maps to automatically detect and classify ASD. 300 ASD fixation maps and 300 TD fixation maps are used for classification. Each image is converted to grayscale and resized with 224 X 224 dimension for processing.

A. Classification of ASD using CNN Training and Testing Phase

The detected section of ultrasound is scaled to 224 x 224 height and breadth for training purposes. The initial convolution for the input image is made up of 32 3X3 filters, yielding 2432 parameters. The activation function for this convolution is ReLU [3]. The convolution layer's 2-dimensional vector is max pooled with a 2X2 matrix [3]. The 2nd convolution of the generated data employs 64 filters with a kernel size of [3] 3X3 array [3], ReLU activation, and yields 51264 parameters [2]. For this convolution, the max pooling filter size is 2X2 [2]. The second convolution produces 51264 parameters. These 51264 data are converted to a single dimension vector using a flattened layer. The first dense layer uses 51380480 parameters to execute nonlinear operations on the flattened input vector. The output of the second dense layer is 257 params. Due to binary level disease classification, sigmoid activation is used in this study; it processes single dimension data as input from a flattened layer [2]. From the input images, this architecture generates 257 features using metrics, loss, and learning rate algorithms. Fit complied data from architecture is utilized to train the model [12]. The training and validation accuracy is 74.33% and 75.23% respectively which is shown in Table I.

TABLE I. CNN TRAINING IMPLEMENTATION WITH OUTPUT PARAMETERS USING PYTHON

Training Accuracy	Validation Accuracy
74.33%	75.23%

Two categories of images with sizes of 224X224 have been added to the batch for testing. The corresponding classes are recognised by the prediction matrix which is given by:

Prediction Matrix =

148	25
20	168

Performance Analysis**:** The confusion matrix [2] is used to calculate the performance analysis for the estimated matrix [3] which is shown on Table II. True positive, true negative, false positive, and false negative [14] scores for the confusion matrix [14] are obtained from the prediction matrix on the ASD validation [2]. Table III represents the performance analysis for classification of ASD on training and validation data. Table IV depicts the performance analysis for classification Accuracy on testing data. The performance methods are designed by using below mentioned formulae:

Accuracy = TP+TN / (TP+TN+FP+FN) [9][10]

Precision = TP/(TP+FP) [9][10]

Recall = TP/(TP+FN) [9][10][21]

F score =2*Precision*Recall/(Precision + Recall)[9][10][21]

TABLE II. CONFUSION MATRIX

	Training samples	
Testing samples	148	25
	20	168

TABLE III. PERFORMANCE ANALYSIS FOR CLASSIFICATION OF AUTISM SPECTRUM DISORDER (ASD) ON TRAINING AND VALIDATION DATA

Layer (type)	Output Shape	Param #
Conv2d (Conv2D)	(None, 224, 224, 32)	2432
max_pooling2d (MaxPooling2D)	(None, 112, 112, 32)	0
dropout (Dropout)	(None, 112, 112, 32)	0
Conv2d_1 (Conv2D)	(None, 112, 112, 64)	51264
max_pooling2d_1 (MaxPooling2D)	(None, 56, 56, 64)	0
dropout_1 (Dropout)	(None, 56, 56, 64)	0
flatten (Flatten)	(None, 200704)	0
dense(Dense)	(None, 256)	51380480
dropout_2(Dropout)	(None, 256)	0
dense_1(Dense)	(None, 1)	257
Total params: 51,434,433 Trainable params: 51,434,433 Non-Trainable params: 0		

TABLE IV. PERFORMANCE ANALYSIS FOR CLASSIFICATION ACCURACY ON TEST DATA

FN	FP	TP	Precision	Accuracy	Recall	F1
0.00	0.00	93.00	0.73	0.75	0.75	0.74
1.00	0.00	92.00	0.73	0.75	0.75	0.74
1.00	1.00	91.00	0.73	0.75	0.75	0.74
2.00	0.00	91.00	0.73	0.75	0.75	0.74
2.00	1.00	90.00	0.73	0.75	0.75	0.74
2.00	2.00	89.00	0.73	0.75	0.75	0.74
3.00	0.00	90.00	0.73	0.75	0.75	0.74
3.00	1.00	89.00	0.73	0.75	0.75	0.74
3.00	2.00	88.00	0.73	0.75	0.75	0.74
3.00	3.00	87.00	0.73	0.75	0.75	0.74
4.00	0.00	89.00	0.73	0.75	0.75	0.74
4.00	1.00	88.00	0.73	0.75	0.75	0.74
4.00	2.00	87.00	0.73	0.75	0.75	0.74
4.00	3.00	86.00	0.73	0.75	0.75	0.74
4.00	4.00	85.00	0.73	0.75	0.75	0.74

FN: False Negative
FP: False Positive
TP: True Positive

The above Tables II, III and IV depicts confusion matrix, performance metrics for classification of ASD using CNN and Performance Analysis for Classification Accuracy on test data, respectively.

Loss function: It is used to calculate model error. The training and validation loss is depicted in Fig. 8. While training neural networks and machine learning models in general [9], thorough going Possibility delivers a structure for selecting a loss function [9]. While training neural network models, the two primary types of loss functions to use are cross-entropy [9] and mean squared error [7][9].

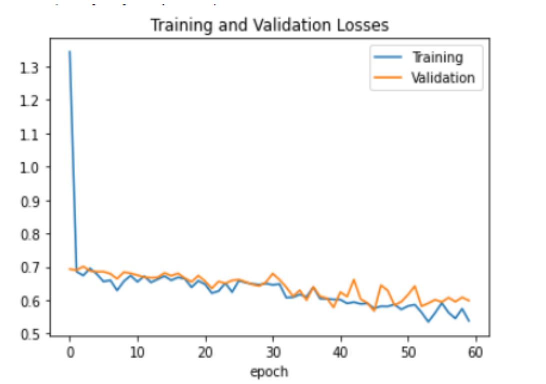


Fig. 8. Training and Validation Loss.

Trainig and validation accuracy: It is the metric for evaluating classification models. Fig. 9 represents the training and validation accuracy.

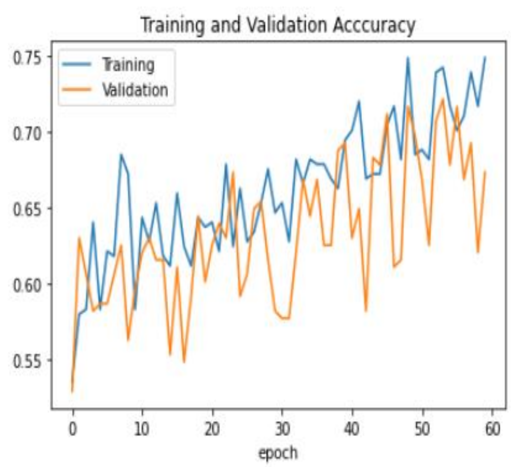


Fig. 9. Training and Validation Accuracy.

X. Conclusion

The major goal of this paper is to classify ASD based on observer's Fixation maps is proposed. The CNN architecture is used to extract the features from fixation maps. 300 ASD fixation maps and 300 TD fixation maps are used for classification. Each image is converted to grayscale and resized with 224 X 224 dimension for processing. The proposed model achieves 75.23% accuracy on the validation dataset.

The objective of this paper is to observe whether eye-tracking data of fixation map could classify children with ASD and typical development (TD). We further investigated whether features on visual fixation would attain better classification performance. In future work, the perceptive technique would be examined and incorporated with the model to enhance the accuracy and other metrics. And also we will try to examine different features of images and can be integrated in our future model. A comprehensive analysis on the dissimilarities in the fixation maps among TD and ASD children will also be accompanied.

Acknowledgment

I would like to express my deep gratitude to my professor Dr. Mahalakshmi R, my research supervisor, for her patient guidance, enthusiastic encouragement and useful critiques of this research work. Finally, I wish to thank my parents for their support and encouragement throughout my study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) declared no funding.

Community Involvement

It is acknowledge that community involvement is not applicable/required.

References

[1] Gutiérrez, Jesús, Zhaohui Che, Guangtao Zhai, and Patrick Le Callet. "Saliency4ASD: Challenge, dataset and tools for visual attention modeling for autism spectrum disorder." Signal Processing: Image Communication 92 (2021): 116092.

[2] Duan, Huiyu, Guangtao Zhai, Xiongkuo Min, Zhaohui Che, Yi Fang, Xiaokang Yang, Jesús Gutiérrez, and Patrick Le Callet. "A dataset of eye movements for the children with autism spectrum disorder." In Proceedings of the 10th ACM Multimedia Systems Conference, pp. 255-260. 2019.

[3] H. Duan, G. Zhai, X. Min, Y. Fang, Z. Che, X. Yang, C. Zhi, H. Yang N. Liu, Learning to predict where the children with ASD look, in: 2018 25th IEEE International Conference on Image Processing (ICIP), 2018, pp. 704–708.

[4] A. Borji, L. Itti, State-of-the-Art in visual attention modeling, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (1) (2013) 185–207. doi:10.1109/TPAMI.2012.89.

[5] Condy, Emma E et al., "Maternal Affect during a Challenging Mother–Child Interaction: The impacts of Broad Autism Phenotype and Respiratory Sinus Arrhythmia Reactivity in Mothers of Children with and Without Autism Spectrum Disorder." Journal of autism and developmental disorders 49, no. 12 (2019): 4891-4900.

[6] Ibrahim, Izaida. "Specific Learning Disorder in Children with Autism Spectrum Disorder: Current Issues and Future Implications." Advances in Neurodevelopmental Disorders (2019): 1-10.

[7] Duan, Huiyu, Xiongkuo Min, Yi Fang, Lei Fan, Xiaokang Yang, and Guangtao Zhai. "Visual attention analysis and prediction on human faces for children with autism spectrum disorder." ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 15, no. 3s (2019): 1-23.

[8] Y. Fang, H. Duan, F. Shi, X. Min, G. Zhai, Identifying children with autism spectrum disorder based on gaze-following, in: 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 423 427.

[9] Le Meur, Olivier, Alexis Nebout, Myriam Cherel, and Elise Etchamendy. "From Kanner Austim to Asperger Syndromes, the Difficult Task to Predict Where ASD People Look at." IEEE Access 8 (2020): 162132-162140.

[10] Tilke Judd, Krista Ehinger, Frédo Durand, and Antonio Torralba. 2009. Learning to predict where humans look. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).2106–2113.

[11] American Psychiatric Association et al. 2013. Diagnostic and statistical manual of mental disorders (DSM-5). American Psychiatric Pub.

[12] W. Jones, K. Carr, A. Klin, "Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder", Archives of general psychiatry (2008) 946–954.

[13] A. Klin, W. Jones, R. Schultz, F. Volkmar, D. Cohen, "Visual fixation pat-terns during viewing of naturalistic social situations as predictors of social competence in individuals with autism", Archives of general psychiatry (2002) 809–816.

[14] Z. Bylinskii, T. Judd, A. Oliva, A. Torralba, F. Durand, What Do Different Evaluation Metrics Tell Us About Saliency Models?, IEEE Transactions on Pattern Analysis and Machine Intelligence 41 (3) (2019) 740–757.doi:10.1109/TPAMI.2018.2815601.

[15] Peter C. Pantelis and Daniel P. Kennedy, "Deconstructing atypical eye gaze perception in autism spectrum disorder", Scientific reports, vol. 7, no. 14990, pp. 1–10, 2017.

[16] Startsev, Mikhail, and Michael Dorr. "Classifying autism spectrum disorder based on scanpaths and saliency." In 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), pp. 633-636. IEEE, 2019.

[17] Tao, Yudong, and Mei-Ling Shyu. "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, 2019.

[18] Rahman, Shafin, Sejuti Rahman, Omar Shahid, Md Tahmeed Abdullah, and Jubair Ahmed Sourov. "Classifying Eye-Tracking Data Using Saliency Maps." In 2020 25th International Conference on Pattern Recognition (ICPR), pp. 9288-9295. IEEE, 2021.

[19] Wang, Wenguan, and Jianbing Shen. "Deep visual attention prediction." IEEE Transactions on Image Processing 27, no. 5 (2017): 2368-2378.

[20] Elbattah, Mahmoud, Romuald Carette, Gilles Dequen, Jean-Luc Guérin, and Federica Cilia. "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, 2019.

[21] Startsev, Mikhail, and Michael Dorr. "Classifying autism spectrum disorder based on scanpaths and saliency." In 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), pp. 633-636. IEEE, 2019.

[22] Ahmed, Zeyad AT, and Mukti E. Jadhav. "Convolutional Neural Network for Prediction of Autism based on Eye-tracking Scanpaths." International Journal of Psychosocial Rehabilitation 24, no. 05 (2020).

[23] Wu, Chongruo, Sidrah Liaqat, Sen-ching Cheung, Chen-Nee Chuah, and Sally Ozonoff. "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, 2019.

[24] Arru, Giuliano, Pramit Mazumdar, and Federica Battisti. "Exploiting visual behaviour for autism spectrum disorder identification." In 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), pp. 637-640. IEEE, 2019.

[25] Goldberg, Joseph H., and Jonathan I. Helfman. "Visual scanpath representation." In Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, pp. 203-210. 2010.

[26] Praveena K.N., Mahalakshmi R. (2022) A Survey on Early Prediction of Autism Spectrum Disorder Using Supervised Machine Learning Methods. In: Rana N.K., Shah A.A., Iqbal R., Khanzode V. (eds) Technology Enabled Ergonomic Design. Design Science and Innovation.Springer,Singapore.https://doi.org/10.1007/978-981-16-6982-8_2.