2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) | 978-1-6654-2642-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/I-SMAC52330.2021.9641046

A Deep Convolutional Neural Network based Detection System for Autism Spectrum Disorder in Facial images

[1]Sajeev Ram Arumugam
Department of AI&DS
Sri Krishna College of Engineering and Technology
Coimbatore, India
imsajeev@gmail.com

[2]Sankar Ganesh Karuppasamy
Department of CSE
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Chennai, India
prof.sankarganesh@gmail.com

[3]Sheela Gowr
Department of CSE
Vels Institute of Science Technology and Advanced Studies(VISTAS)
Chennai, India
sheela.se@velsuniv.ac.in

[4]Oswalt Manoj
Department of CSBS
Sri Krishna College of Engineering and Technology
Coimbatore, India
oswaltmanoj@skcet.ac.in

[5]K Kalaivani
Department of CSE
Vels Institute of Science Technology and Advanced Studies(VISTAS)
Chennai, India
kalai.se@velsuniv.ac.in

***Abstract*— Autism spectrum disorder (ASD) a development disability which causes several challenges in social, communication and behaviour abilities. ASD people are nowhere much different when compared to ordinary people rather the way they interact will be different, few ppl of ASD needed help for all basic needs and others don't. On effectively identifying ASD at the earlier stages helps to provide therapy to improve their skills. Being a disability in neurological development, many researchers are trying to predict the ASD in advance with Image processing techniques based on MRI Images. This research work has attempted to develop a prediction system based on Convolution Neural Network [CNN] based on their photos. Database for the required model is taken from Kaggle and split into 80:20 for training and testing the model. Our model managed to give an accuracy rate of 91% and a overall loss of 0.53.**

Keywords*— Autism Spectrum Disorder(ASD), Autism detection, Face recognition, Neural Network, Functional magnetic resonance imaging(fMRI), CNN, Centres for Disease Control and Prevention (CDC).

I. Introduction

According to the Centers for Disease Control and Prevention, one in fifty-four children gets affected due to ASD [1]. ASD affects four times further boys than girls, and ASD diagnoses one in six children aged three to seventeen. Every child with ASD will not have the same symptoms, and each child differs. Some common symptoms are lack of eye contact, not pointing to objects, not showing anxiety for others, making friends, not communicating, toe walking, flapping hands, doesn't seem to feel pain, etc. The count of autism in India is increasing over the years. It may be due to earlier detection, diagnostic substitution etc. [2]. The reason for ASD is due to genetic and environmental factors.

Parents and paediatricians perceptions play a significant role in the early prediction of ASD. If the children identified with ASD was not treated in the early stages, they have poor academics, hyperactivity and aggression. It leads to anxiety and depressions. There are many standardized tools for the prediction of ASD, such as Autism Spectrum Quotient (ASQ), Social Communication Disorder Checklist (SCDC), Modified Checklist for Autism in Toddlers (M-CHAT), Social Responsiveness Scale (SRS), Social Communication Questionnaire (SCQ) and Autism Behaviour Checklist (ABC) [3].

Early detection and treatment are the primary steps to normalize the kid to social and behavioural attitude. There is no medical test available for the prediction of autism. The diagnosis of ASD takes a long time, and the increase in ASD cases worldwide is a motivation for doctors/scientists to find more effective screening methods. In India, around 30 lakhs of children need therapists regularly and are not practically feasible, due to availability of large number of therapy centres and therapist and hence parental therapy becomes mandatory [4]. Parents can also be an effective co-therapists and as adequate substitute intervention for children with ASD. In some government centres, the parents are trained to provide therapy to their child daily.

II. Literature Survey

Xiaoxiao [5] applied a deep neural network in fMRI images to identify ASD and control subjects. DNN classifier, along with sampling and significance testing scheme, produces reliable results. The classifier achieves 85.3% accuracy. Saad Sadiq [6] used a parallel approach which examine language modalities and determined relations between objective measurements of social, behavioural, communication skills and ASD symptoms. Here, Long Short Term Memory (LSTM) networks were used for Speech Activity Detection. The system was able to achieve F-Static value of 20.84.

Matthew [7] used a random forest classifier to predict ASD in children. The dataset was collected in Metropolitan Atlanta Developmental Disabilities Surveillance Program. The first step is to find the group of words and phrases that are mandatorily needed for ASD classification. The second step is to build an algorithm to perform the classification. The system achieves 84.0% sensitivity and 89.4% positive predictive value. Wenbo Liu [8] used SVM Classification using face images to classify the children with ASD. The algorithm method analyses the eye movement to classify children with and without ASD. The accuracy of the machine learning model was identified to be 88.51%.

Kazi Shahrukh Omar [9] developed a mobile application based prediction model to identify ASD for children. The mobile based prediction model was developed by parallelising Random Forest-CART and Random Forest. The proposed system was developed taking the help of AQ10 dataset and 250 real datasets collected from people with and without ASD which was used for training and testing. The system managed to give an accuracy of 92.26%. Xia-an [10] uses Support Vector machine to classify ASD patients. In the SVM based method, fMRI data of 61 ASD patients and 46 TC was obtained from the Autism Brain Imaging Data Exchange (ABIDE) database. The SVM based machine was developed and used for classifying ASD from TC. The accuracy of the SVM based classification system was found to be 96.15%.

Wasifa [11] used 12 subjects of ASD and non-ASD patients EEG for finding fearful face, happy face and neutral face. The minimal and maximal values for each subject are used to extract brain connectivity features for classification amongst ASD and non-ASD. The discriminant analysis and support vector machine with polynomial kernels are used for classification. The classification algorithm attains 94.7% accuracy, 85.7% sensitivity and 100% during cross-validation. Yazhou Kong [12] used conventional methods to classify ASD and TC by extracting morphological features at diverse regions of interest and connectivity between ROI[13]. A deep neural network classifier is used to predict between Autism Spectrum Disorder and non-Autistic Disorder. The proposed method attains an accuracy of 90.39%.

Lingyu Xu [14] proposed a method which evaluates the global time-varying behaviour of brain activity by calculating the change in first-order statistical properties from the functional near-infrared spectroscopy time series. A deep learning model was constructed to find the potential patterns of temporal variation for identifying Autism Spectrum Disorder. The system achieved a classification accuracy of 95.7%. Romuald Carette [15] used eye tracking scan paths for binary image classification. The classifier predicts participants into two categories. The system achieves 90% accuracy using an ANN classifier.

The performance metrics for various models are discussed in Tab.1.

III. Proposed System

The Computer-Aided Detection system has helped in the medical field in diagnosing cancer, blood clots etc. The proposed Computer-Aided Detection provides a diagnostic tool that helps parents decide to pursue further ASD testing. We have used the autism dataset taken from Kaggle in the proposed system, which has 2940 autistic and non-autistic images. The dataset contains an age distribution of 2 to 14 years [16]. The gender ratios are near to their corresponding populations. Men are diagnosed with autism three times more than women. In the not autistic class, the balance is closer to 1:1 [17]. The system architecture of the ASD detection system is shown in Fig. *1*

Tab.1. Performance metrics of different models

Author	Database	Methods	Performance Measures
Xiaoxiao et al.[5]	ABIDE	DNN Classifier	Accuracy - 85.3%
Saad Sadiq et al. [6]	clinical data - 33 children	LSTM Network	F-Static value- 20.84.
Matthew et al. [7]	ADDM site	Random Forest Classifier	Sensitivity -84% Positive predictive value -89.4%
Wenbo Liu et al. [8]	Clinical data -87 children	SVM classificatio n	Accuracy - 88.51%
Kazi Shahrukh Omar et al. [9]	250 real datasets	Random Forest Classifier	Accuracy - 92.26%
Xia-an et al. [10]	ABIDE	SVM classificatio n	Accuracy - 96.15%
Wasifa et al. [11]	Clinical data- 12 subjects	Discrimina nt analysis and support vector machine	Accuracy - 94.7% Sensitivity - 85.7%
Yazhou Kong et al.[12]	ABIDE	Deep Neural Network	Accuracy - 90.39%
Lingyu Xu et al.[14]	Clinical data -22 children	LSTM and CNN	Accuracy - 95.7%
Romuald Carette et al. [15]	Clinical data -59 children	ANN classifier	Accuracy -90%

Convolution Neural Network is used to detect critical points on face images. The Kaggle dataset is used as a training dataset for autism identification and verification. The VGG face model trains on a large dataset and evaluates autism datasets, demonstrating that the model effectively generates features from faces [18]. It describes the process of training an autism detection classifier in which softmax activation function is used in the output layer to classify autism faces [19].

The neural network get the image as input, later assign default weight and bias to the various aspects of the image and classifies it into children with and without ASD. as we are going to use convolution neural network the pre-processing of image can be avoided as it holds filters inbuild to itself [20]. The convolution neural network is very similar to the architecture of human brain. The spatial and temporal dependencies of an image is captured by applying relevant filters in a convolution neural network. The architecture performs an enhanced fitting owing to the decline in the number of parameters involved and the weights reusability [21].

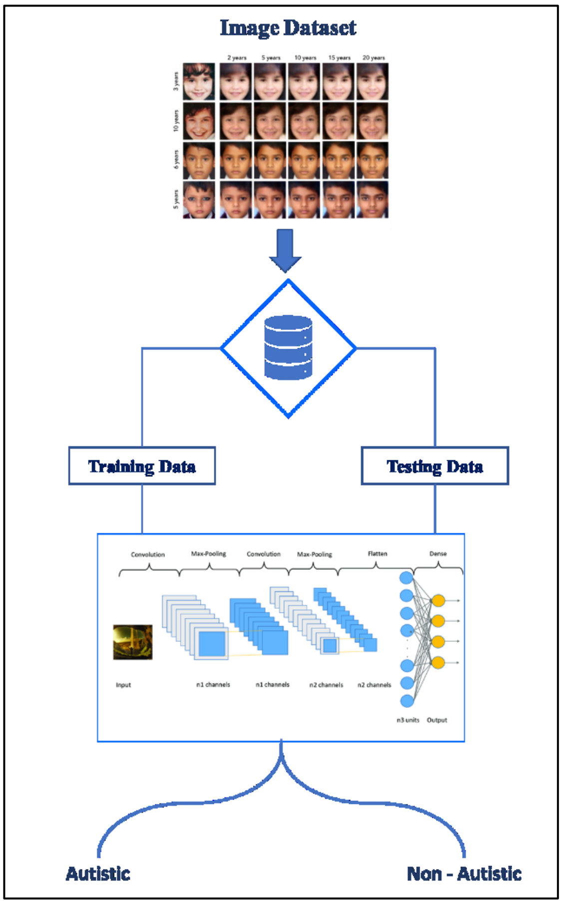

Fig. 1 The architecture of the proposed ASD detection

The neural network get the image as input, later assign default weight and bias to the various aspects of the image and classifies it into children with and without ASD. as we are going to use convolution neural network the pre-processing of image can be avoided as it holds filters inbuild to itself [20]. The convolution neural network is very similar to the architecture of human brain. The spatial and temporal dependencies of an image is captured by applying relevant filters in a convolution neural network. The architecture performs an enhanced fitting owing to the decline in the number of parameters involved and the weights reusability [21].

The designed Conventional Neural Network contains sequence of convolution, maxpool layers. The input layer gets the input image and collects the basic information such as colour, edge and gradient orientation. The follow up layers in the system collects the high level features [22]. Max Pooling layer discards the noise available in the images and parallelly performs dimensionality reduction. depending on the type of image, the number of convolution layers can be added for gathering low-level information.

A Fully-Connected convolutional layer learns the combination of the higher order features. The input image is flattened to column vectors and then fed to a feed-forward neural network, and enhanced for every training iteration [23]. After many epochs, the model differentiated between high-level and low-level features in images and classifying them using the neural network. Fig. 2 shows the visualization of the proposed system.

The autism dataset is used as an input image with 2940 autistic and non-autistic images in the proposed system. Pairs of 2D convolution layer and max pool layers are used to create a convolution kernel starting from the input layer to the output layer to produce an output tensor. The first convolution layer is in size of 300 X 300 matrix. The filter specifies the number of output filters in the convolution. In the first layer, we used 64 filters. The height as well as width of the convolution window is specified by Kernel size. In the first layer, the height of the convolution window is 3, and the width of the convolution window is 3. Padding indicates that output has the same height/width dimension as the input. ReLU activation allows the model to learn and performs better.

In the second convolution layer,64 output filters are used. The convolution window is set to 3X3. Padding is used to have the same height and weight as the input and ReLU activation function. Max Pooling calculates the maximum value for each patch of the image. They are helpful when the convolutional layer detects small changes in the feature location. Pool size takes the maximum value of a 2X2 pooling window. Strides values specify the movement of pooling window for each pooling step. In the third convolution layer and fourth convolution layer, 128 output filters are used. The convolution window is set to 3X3. Padding has the same height and weight as the input and ReLU activation function.

In the second max pooling, the Pool size takes the maximum value of a 2X2 pooling window. Strides values specify the movement of pooling window for each pooling step, and it is set to 2X2. In the fifth convolution layer and sixth convolution layer, 256 output filters are used. The convolution window is set to 3X3. The exact height and weight of padding is used as the input and activation functions set ot ReLU. In the third max pooling, the maximum value of a pooling window is 2X2. Strides specify the pooling window moves for each pooling step. In the seventh convolution layer and eighth convolution layer, 512 output filters are used. The convolution window is set to 3X3. The exact height and weight of padding are used as the input and activation functions set ot ReLU.

In the fourth max pooling, the maximum value of a pooling window is 2X2. Strides values specify the movement of pooling window for each pooling step. In the ninth convolution layer and tenth convolution layer, 512 output filters are used. The convolution window is set to 3X3. The same height and weight of padding are used as the input and activation functions set ot ReLU. In the fifth max pooling, the maximum value of a pooling window is 2X2. Strides specify the pooling window moves for each pooling step. Flatten is used to flatten the input. It is used to preserve weight when switching from one data format to another format. After creating convolution, the data is passed to the dense layer with a dense layer consisting of 4096 units and a dense Softmax layer of 2 units to avoid forwarding negative values in the network. **Error! Reference source not found.** shows the summary of the classification system.

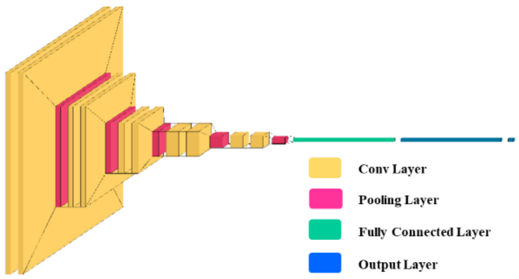

Fig. 2. Visualization of the proposed Neural Network

Layer (type)	Output Shape	Param #
conv2d_26 (Conv2D)	(None, 300, 300, 64)	1792
conv2d_27 (Conv2D)	(None, 300, 300, 64)	36928
max_pooling2d_10 (MaxPooling	(None, 150, 150, 64)	0
conv2d_28 (Conv2D)	(None, 150, 150, 128)	73856
conv2d_29 (Conv2D)	(None, 150, 150, 128)	147584
max_pooling2d_11 (MaxPooling	(None, 75, 75, 128)	0
conv2d_30 (Conv2D)	(None, 75, 75, 256)	295168
conv2d_31 (Conv2D)	(None, 75, 75, 256)	590080
max_pooling2d_12 (MaxPooling	(None, 37, 37, 256)	0
conv2d_32 (Conv2D)	(None, 37, 37, 512)	1180160
conv2d_33 (Conv2D)	(None, 37, 37, 512)	2359808
max_pooling2d_13 (MaxPooling	(None, 18, 18, 512)	0
conv2d_34 (Conv2D)	(None, 18, 18, 512)	2359808
conv2d_35 (Conv2D)	(None, 18, 18, 512)	2359808
max_pooling2d_14 (MaxPooling	(None, 9, 9, 512)	0

Fig 3. CNN for the proposed system

IV. Experimental Results

The learning curve diagnoses performance of the training dataset and testing dataset. When it comes to evaluating models, separating data into training and testing sets is crucial. When we divide a data set into a training set and a testing set, we use the majority of the data for training and a smaller piece for testing. We can reduce the consequences of data inconsistencies and better understand the model's properties by using similar data for training and testing. We test a model by generating predictions against the test set after it has been processed using the training set. Here, we used 80% of the data for training and 20% of the data for testing. The epochs define the number of times the developed model train through the complete training dataset. Every epoch updates the internal model parameters for every sample in the dataset. Figure 4 shows the accuracy of the training and testing data used. For every epoch, the accuracy increased for training as well as testing dataset. When epoch reaches 70, the system accuracy is 91%, and then the system starts to overfit. The accuracy of the system is calculated using the formula: The loss for the proposed training and the testing dataset is shown in Fig. 5. Both the training and testing dataset has a loss of 3 at 10 epochs, and it starts to decrease and reaches 0.53 loss when it reaches 70 epochs.

The performance metric accuracy is the ratio of appropriately examined to all examined as shown in Equation (1).

$$Accuracy = \frac{Correctly\ predicted}{Total\ number\ of\ predictions} \ldots\ldots\ldots(1)$$

The error for the present model state must be estimated several times. This necessitates the selection of an error function, referred to as a loss function, that may be used to estimate the model's loss and update the weights to lower the loss on the next assessment. The Mean Square Error is the sum of squared distances between our target variable and predicted values as shown in Equation (2).

$$MSE = \frac{1}{N}\sum_{i=0}^{N}(y - \hat{y_i})^2 \ldots\ldots\ldots(2)$$

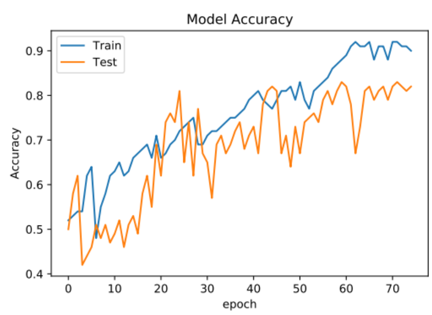

Fig. 3 Accuracy for every epoch in training and testing dataset

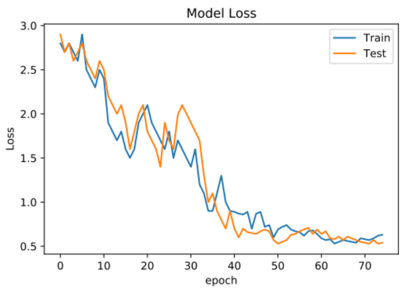

Fig. 4 Loss for every epoch in training and testing dataset

V. Conclusion

This paper presents an Autism Spectrum Detection system based on CNN architecture. the system was initially developed using machine learning techniques and later modified for neural networks. The System performed well for Convolution Neural Networks and later enhanced to improve the classification results. The layer counts, weights were manually adjusted so that the system was not under or over performed. we have used accuracy and loss value evaluation parameters to identify the performance of the system which was identified to be 91% and 0.53 respectively. The system could be further developed to identify the stages of ASD if real time datasets are collected and trained accordingly which helps further to identify the stage of ASD.

References

[1] "Centers for Disease Control and Prevention." [Online]. Available: https://www.cdc.gov/.

[2] "Data & Statistics on Autism Spectrum Disorder | CDC." [Online]. Available: https://www.cdc.gov/ncbddd/autism/data.html.

[3] E. E. Bolte and J. J. Diehl, "Measurement tools and target symptoms/skills used to assess treatment response for individuals with autism spectrum disorder," *J. Autism Dev. Disord.*, vol. 43, no. 11, pp. 2491–2501, 2013.

[4] K. Bearss, T. L. Burrell, L. Stewart, and L. Scahill, "Parent Training in Autism Spectrum Disorder: What's in a Name?," *Clin. Child Fam. Psychol. Rev.*, vol. 18, no. 2, pp. 170–182, 2015.

[5] X. Li, N. C. Dvornek, J. Zhuang, P. Ventola, and J. S. Duncan, *Brain biomarker interpretation in ASD using deep learning and fMRI*, vol. 1. Springer International Publishing, 2018.

[6] S. Sadiq, M. Castellanos, J. Moffitt, M. L. Shyu, L. Perry, and D. Messinger, "Deep learning based multimedia data mining for autism spectrum disorder (ASD) diagnosis," *IEEE Int. Conf. Data Min. Work. ICDMW*, vol. 2019-Novem, pp. 847–854, 2019.

[7] M. J. Maenner, M. Yeargin-Allsopp, K. N. Van Braun, D. L. Christensen, and L. A. Schieve, "Development of a machine learning algorithm for the surveillance of autism spectrum disorder," *PLoS One*, vol. 11, no. 12, pp. 1–11, 2016.

[8] W. Liu, M. Li, and L. Yi, "Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework," *Autism Res.*, vol. 9, no. 8, pp. 888–898, 2016.

[9] K. S. Oma, P. Mondal, N. S. Khan, M. R. K. Rizvi, and M. N. Islam, "A Machine Learning Approach to Predict Autism Spectrum Disorder," *2nd Int. Conf. Electr. Comput. Commun. Eng. ECCE 2019*, pp. 7–9, 2019.

[10] X. Bi, Y. Wang, Q. Shu, Q. Sun, and Q. Xu, "Classification of Autism Spectrum Disorder Using Random Support Vector Machine Cluster," *Front. Genet.*, vol. 9, no. FEB, p. 18, Feb. 2018.

[11] W. Jamal, S. Das, and I. Oprescu, "Classi fi cation of autism spectrum disorder using supervised learning of brain connectivity measures extracted from synchrostates," *J. Neural Eng.*, vol. 046019, 2014.

[12] Y. Kong, J. Gao, Y. Xu, Y. Pan, J. Wang, and J. Liu, "Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier," *Neurocomputing*, vol. 324, pp. 63–68, 2019.

[13] S. Ram, K. Kalaivani, A. Sahayadhas, and C. S. Shylaja, "Intensified computer aided detection system for lung cancer detection using MEM algorithm," *Int. J. Pharm. Res.*, vol. 12, no. 2, pp. 643–647, 2020.

[14] L. Xu *et al.*, "Characterizing autism spectrum disorder by deep learning spontaneous brain activity from functional near-infrared spectroscopy," *J. Neurosci. Methods*, vol. 331, p. 108538, 2020.

[15] R. Carette, M. Elbattah, F. Cilia, G. Dequen, J. L. Guérin, and J. Bosche, "Learning to predict autism spectrum disorder based on the visual patterns of eye-tracking scanpaths," *Heal. 2019 - 12th Int. Conf. Heal. Informatics, Proceedings; Part 12th Int. Jt. Conf. Biomed. Eng. Syst. Technol. BIOSTEC 2019*, pp. 103–112, 2019.

[16] "Detect Autism from a facial image | Kaggle." [Online]. Available: https://www.kaggle.com/gpiosenka/autistic-children-data-set-traintestvalidate/version/5.

[17] R. Loomes, L. Hull, and W. P. L. Mandy, "What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis," *J. Am. Acad. Child Adolesc. Psychiatry*, vol. 56, no. 6, pp. 466–474, 2017.

[18] R. Vinayakumar, K. P. Soman, and P. Poornachandrany, "Applying convolutional neural network for network intrusion detection," *2017 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2017*, vol. 2017-Janua, pp. 1222–1228, 2017.

[19] S. Albawi, T. A. M. Mohammed, and S. Alzawi, "Layers of a Convolutional Neural Network," *Ieee*, 2017.

[20] Z. Zhang, "Derivation of Backpropagation in Convolutional Neural Network (CNN)," *Univ. Tennessee, Knoxville, TN*, pp. 1–7, 2016.

[21] R. Chauhan, "Convolutional Neural Network (CNN) for Image Detection and Recognition," *2018 First Int. Conf. Secur. Cyber Comput. Commun.*, pp. 278–282, 2018.

[22] R. Yamashita, M. Nishio, R. Kinh, G. Do, and K. Togashi, "Convolutional neural networks : an overview and application in radiology," *Insights Imaging*, 2018.

[23] Y. Demir, "Face Recognition Based on Convolutional Neural Network," *Int. Conf. Mod. Electr. Energy Sytems*, pp. 376–379, 2017.