International Journal of Future Generation Communication and Networking
Vol. 13, No. 4, (2020), pp. 2432–2444

A Convolutional Neural Network Approach To Identify Presence Of Autism In Facial Images

*Amrita Budarapu[1], Dr. Nara Kalyani[2]

[1]*Department of Computer Science, GNITS, Telangana, Hyderabad,*
[2] *Professor, Department of Computer Science, GNITS, Telangana, Hyderabad,*
[1]*amritabudarapu18@gmail.com*
[2]*nara.kalyani@gnits.ac.in*

Abstract

ASD (Autism Spectrum Disorder) is characterized as a neurological developmental disorder by the reduced ability of the children which affects their social interaction, language (or) behavioral skills. This influences to engage in repetitive and stereotypic behaviors. The underlying causes of ASD are still not well understood and diagnosed, but it is increasing at an alarming rate and the number of children suffering from this disorder are also rapidly increasing. Facial appearance (or) expression is the most powerful and natural non-verbal emotional communication method. They are very important in our day to day conversation, next to the tone of voice. But the kids with Autism Disorder may have the disability to communicate properly. Several existing systems aids in knowing the severity of the disorder and the child's ability to present the facial expressions, to the verbal prompt provided. However, these systems work on the children having ASD. But there is a gap in identifying the existence of autism in the children at early stages. The proposed model addresses this issue based on facial expressions. The model detects the facial landmarks from the input image and extracts the key ROI (Region of Interest). Detection of facial landmarks from the input image is a part of the shape prediction issue. Given an input image, a shape predictor tries to localize the key points of interest (x, y coordinates) along the shape. These key points (features) are given as input to the CNN (Convolutional Neural Network) model and are trained to detect the presence of Autism in Kids at early stages. For experimental purpose Autistic Children Dataset available in Kaggle is used, and the implemented model is validated using internal validation as well as external validation with an accuracy of 81% and 73% respectively.

Keywords— *Autism Spectrum Disorder (ASD), Convolutional Neural Network (CNN), Facial landmarks.*

I. INTRODUCTION

Human beings generally express their emotions in various ways, such as facial expressions, bodily expressions, or through language. ASD (Autism Spectrum Disorder) is a neuro developmental disorder that is characterized by, and diagnosed on the grounds of, social communication deficits and behavioral deficits. Impairment in social communication is one of the core feature of Autism Spectrum Disorder (ASD). Because the ability to derive any individual's emotions or feelings from their facial expressions is crucial in many aspects of social conversations or gatherings. Deficits in expression recognition are a plausible candidate marker for ASD [19]. Autism is one of the rapidly growing developmental disorders in the world today. It is a life-long neuro developmental disorder, behaviorally well-defined by impairments in social communication and the existence of repetitive and restricted behaviors and interests. For example, social communicative impairments can impact in social-emotional responses, in non- verbal communicative behaviors, and in development and maintenance of social relationships [20].

As the new innovative system is a great platform for assisting the autistic therapies, it also comes with challenges. As an example, it is common for a person with ASD not to glaze their eyes as a sign of their attention in social interactions. This is the main focuses in many intervention therapies since it is a great biofeedback from the user about their social attention. This can be quite challenging [21]. One of the main concerns to deal with this issue is how to gain benefit out of the technology, to help these children suffering from this disorder [5]. Individuals with intellectual disabilities and developmental disorders create another rapidly growing population that may gain from assistive applications in the contexts of special therapy, education and training. Even though most of the socially aided research till date is focused on ASD, to enable a positive long-term outcome early intervention is very curtail, so that the treatment can be done successfully [5].

Objective: The model proposed, predicts the presence of Autism in Kids at early stages based on facial landmarks in the images.

Motivation: Early intervention of the disorder can help the children in getting the treatment with utmost care in many ways, which may help them in long run.

Contribution: The Paper aims in predicting the presence of Autism at early stages. Here in this model an image is taken and, at the pre-processing stage the input image is cropped with ROI (Region of Interest) that indicates the object of interest, and a shape predictor tries to fetch key points of interest from the shape which extracts the facial landmarks in the images. The CNN model is then trained with Autism Dataset which classifies Autistic behavior based on extracted facial landmarks.

The following paper implements a Convolutional Neural Network (CNN) model to Identify presence of Autism in Facial images and is organized by introducing the Autism Spectrum Disorder (ASD) in section I. Section II details the related work of various contributions contributed by various researchers in this area of research. Section III provides insights of the proposed

system wherein we have System Architecture, Pre-processing of the Images and the CNN model. Section IV describes the implementation of the proposed system in detail: the dataset taken, training the model, and testing the model. Results obtained and the Analysis are discussed in Section V. Finally, Section
VI concludes the work done and provides future scope of the project. And the references are listed at the end.

II. RELATED WORK

Autism is one of the rapidly flourishing developmental disorders in the world today. It is a neuropsychiatric disorder characterized by severe and sustained impairment in social interaction, deviance in communication, and patterns of behavior that are restricted, stereotyped, or both [10]. Autism spectrum disorder is a general term for a group of complex disorders of brain development. These disorders are categorized, in varying degrees, by having problems in social interaction, verbal and nonverbal communication and repetitive behaviors. Both adults and children with ASD characteristically show difficulties in verbal and non-verbal communication, social interactions, and leisure or play activities and these behaviors continue into adulthood [3].

Autism Spectrum Disorder is one among the most puzzling disorders for which no exact cause has been identified but also no exact medication has been found yet. Over the decades, its prevalence showed a dramatic rise; an observation that encouraged many researchers around the globe to try to explore all aspects from etiology to diagnosis and treatment. Although autistic caregivers as well as physicians and researchers would like to know the actual origin of ASD and find a certain cure for it, this goal still seems to be distant. Accordingly, the realistic achievable goal to deal with this disorder is to try to direct all the applicable resources to help its sufferers to improve their skills and

ISSN: 2233-7857 IJFGCN

functioning and get the utmost benefits of their strengths aiming at improving their quality of life. Socio-economic support for autistic caregivers is also essential to empower them in helping their children across their journey for a better tomorrow [5].

Autism Spectrum Disorder (ASD): Journey from the past to the present

Historically, late in the first decade of the twentieth century, Eugen Bleuler used the term autism to describe some schizophrenic symptomatology. That term was derived from the Greek word “autos” that means “self” to reflect the status of a person who has impaired social reciprocity, hence lives isolated. Later, in 1940s, the term “autism” was used to refer to children suffering from social and or emotional problems. By that time, Leo Kanner from USA called several children with tendency to remove themselves from their surrounding societies “autistics”. Concurrently, Hans Asperger from Germany described similar cases suffering from the same problem that later was identified as Asperger syndrome [20,21].

Autism has been identified as a separate disorder from schizophrenia only in 1960s and since then to 1970s it has been treated using electroconvulsive therapy, LSD, and behavioral modification techniques depending on all types of punishment but from 1980s to 1990s till the present; behavior therapy using positive reinforcement, environmental restructuring for learning, and communication skills training have become the main therapeutic modalities for ASD [21-23].

As maximum of the socially assistive research till date concentrates on Autism Spectrum Disorder, early detection is critical to enable a positive long-term outcome, so that it can be treated successfully.

Beibin Li [5] analyzed facial images from various settings like illumination, pose, occlusion etc. and then used machine learning techniques for classifying ASD using facial aspects such as action units, expression, valence, and arousal. This system classifies ASD using representations of various facial attributes from CNN, which are trained on input images from the wild, but with the limited set of training data.

This research aims on the ability of the individuals with this disorder (ASD) to produce facial expressions of emotions with respect to verbal prompt [6]. They used the Janssen Autism Knowledge Engine (JAKE®), including automated facial expression analysis software (FACET) to measure facial expressions in individuals with ASD (n = 144) and a typically developing (TD) comparison group (n = 41). Differences in ability to produce facial expressions were observed between ASD and TD groups, demonstrated by activation of facial action units (happy, scared, surprised, disgusted, but not angry or sad).

This area of research aims on understanding the frequency and severity of impairments in the participants with ASD on the Films Expression Task [7], which combines three key features of real-life expression recognition: naturalistic facial expressions, a broad range of simple and complex emotions, and short presentation time. Findings indicate that most people with ASD have severe expression recognition deficits and that the Films Expression Test is a sensitive task for biomarker research in ASD

This paper introduces an innovative system to recognize facial expressions in children with ASD during playtime [8]. Children are observed while playing or using their tablets or laptops while the researchers track the child's facial expressions. An Active shape Model (ASM) tracker was implemented, which tracks 116 facial landmarks via web-cam input, the tracked landmark points are used to extract face expression features using a Support Vector Machine (SVM) based classifier which gives rise to robust our system by recognizing seven expressions rather than only six expression as in the most of face expression systems. The system is applied to Child Affective Face Expression CAFE set.

A study proposed the use of SVM (support vector machine) to detect autistic qualities in individuals [10]. The underlying data used for this study is taken from the “Autism Brain Imaging Data Exchange” (ABIDE) database known as “Resting state functional magnetic resonance imaging” data (FMRI). The dataset has 107 instances and only 84 instances are used in the actual study because the rest were not up to the specifications of the study. This research stated that the support vector machine achieved a projecting accuracy of 96%; it further concluded that the

ISSN: 2233-7857 IJFGCN

performance is maximized and stabilized when the first 272 features are considered to be the optimal feature set, out of 500 optimal base classifiers.

One of the recent research papers worked on predicting autism in kids by using different classification methods [3]. The dataset used was obtained through the ASDTests app, and the considered age group of children in autism children dataset was between 4-11 years. The dataset had 292 instances and 19 dissimilar attributes. Different models were built using 70% training data and 30% testing data, and the model is evaluated using the performance metrics accuracy, sensitivity, specificity, precision, and recall. The study proved that overall performance increased in terms of all the performance metrics.

III. PROPOSED ARCHITECTURE

The below architecture depicted in figure 1, represents the process flow of the work done which briefs how it was implemented with two phases, the training phase and the testing phase. Taking the sample images, pre-processing the images to extract the features using facial landmarks and training the CNN model comes under the training phase and classification of the images comes under the testing phase.

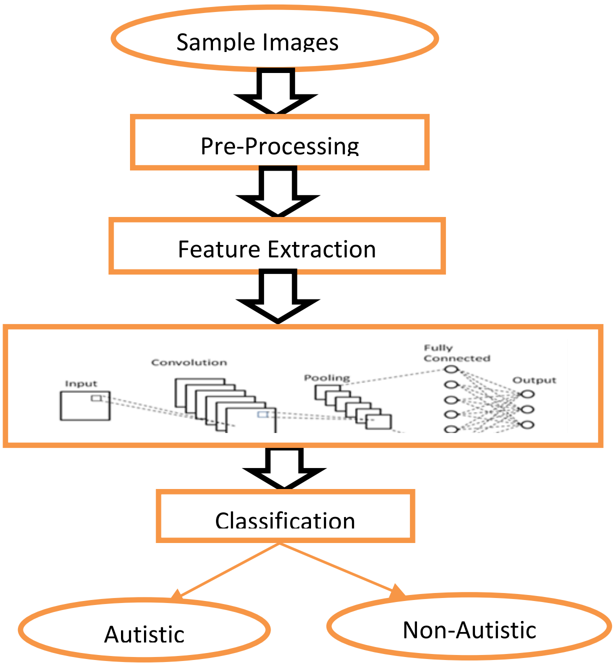


Figure 1: System Architecture

The description of the technology used is described in detail in the following sections.

Pre-processing

Pre-processing is done to extract the Facial landmarks from the image given using shape prediction methods.

Facial Landmarks

Detection of facial landmarks from a given input image is a part of the shape prediction problem. Given an input image normally an ROI (Region of Interest) that specifies the object in the image is considered and a shape predictor tries to extract main points of interest along the face region [26].

Through facial landmarks, the main aim is to detect facial structures (key points) on the face by using some shape prediction methods included in Dlib.

Detecting facial landmarks can be performed in two-steps:

Step #1: Localizing the key points in the face from the image.

Step #2: Detecting the crucial facial structures on the face.

ISSN: 2233-7857 IJFGCN

Step 1: Face detection

By applying a pre-trained **HOG (Histogram of Gradient Descents) and Linear SVM (Support Vector Machine) object detector,** the process of face detection from a given image is done successfully.

HOG and Linear SVM object detector

The Histogram of Oriented Gradients method suggested by Dalal and Triggs in their seminal 2005 paper, "***Histogram of Oriented Gradients for Human Detection"*** demonstrated that the Histogram of Oriented Gradients (HOG) image descriptor and a Linear SVM can be used to train highly accurate object classifiers — or in their particular study, human detectors [31].

HOG is based on the idea that local object appearance can be effectively described by the distribution (histogram) of edge directions (oriented gradients).

General algorithm for training an object detector using Histogram of Oriented Gradients:

1. Gradient calculation: Calculate the x and the y gradient images.
2. Cells: Divide the image into cells (Ex: 3x3).
3. Calculate histogram of gradients in these cells.
4. Block normalization: To counter the effects histogram is normalized.
5. Feature Vector: Final feature vector is calculated from the normalized histogram.

For Classification of the normalized feature SVM (Support Vector Machine) algorithm is used.

The actual algorithm used to detect the face regions in the image doesn't matter. The important part here is that through some method we obtain the bounding box (i.e., (x, y) coordinates of the face in the image).

Given the face region we can then apply step 2,

Step 2: Detecting crucial facial structures on the face. Facial features or regions:

- Mouth
- Right eyebrow
- Left eyebrow
- Right eye
- Left eye
- Nose
- Jaw

The facial landmark detector method is present in the Dlib library. The result of this detector is to detect facial landmarks in real- time with high quality predictions.

Understanding dib's facial landmark detector

The pre-trained facial landmark detector implemented inside the dlib library is used to estimate the location of ***68 (x, y)- coordinates*** that map to facial structures (key points) on the face region [28].

The indexes of the 68 coordinates can be visualized on the image below:

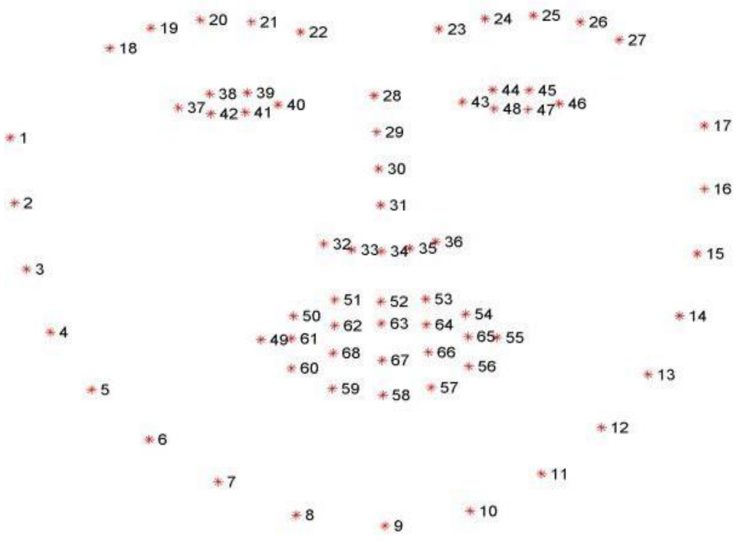


Figure 2. Visualizing the 68 facial landmark coordinates

ISSN: 2233-7857 IJFGCN

Facial regions can be accessed through simple Python indexing (assuming zero-indexing with Python since the image above is one- indexed) The indexes are listed below:

- Mouth: [48, 68].
- Right eyebrow: [17, 22].
- Left eyebrow: [22, 27].
- Right eye: [36, 42].
- Left eye: [42, 48].
- Nose: [27, 35].
- And the Jaw: [0, 17].

Procedure

1. Load face detector: All facial landmark detection algorithms take as input a cropped facial image. Therefore, our first step is to detect all faces in the image and pass those face rectangles to the landmark detector.
2. Load landmark detector: Next, load the landmark detector (shape predictor.dat) This landmark detector was trained on a few thousand images of facial images and corresponding landmarks.
3. Capture frames and process it.
4. Detect faces: Run the face detector on every frame of the Image. The result of a face detector is a vector of rectangles that has one or more faces in the image.
5. Run facial landmark detector: Pass the original image and the detected face rectangles to the facial landmark detector. For every face, gets 68 landmarks which are stored in a vector of points. Because there can be multiple faces in a single frame, which must pass a vector of points to store the landmarks.
6. Draw landmarks: Once it has obtained the landmarks, can draw them on the frame for display.

Convolutional Neural Network

A **Convolutional Neural Network (ConvNet/CNN)** is a Deep Learning technique which takes the given input image, assigns learnable weights and biases to various objects in the image and will differentiate one object from the other. The pre-processing needed in CNN is very less when compared to other classification algorithms. In other primitive methods filters are provided, with required training, through which ConvNets can learn these filters/characteristics [27].

The architecture of CNN is very much similar to that of the connectivity pattern which Neurons have in Human Brain and was influenced by the organization of Visual Cortex. Generally, each neuron responds to stimulus in a restricted area of the visual field known as the Receptive Field. A collection of such fields covers the entire visual area.

Image classification in CNN takes an input image, process it and classifies into certain categories (E.g., Dog, Cat, Tiger, Lion). A Computer considers any image as an array of pixels, and it is based on the image resolution. Image resolution is taken by h x w x d. E.g., A color image of 6 x 6 x 3 array of matrix, where 3 refers to RGB values, an image of 4 x 4 x 1 array of matrix, where 1 refers to grayscale image.

Deep learning techniques like CNN models are used for training and classification of images. Each image will pass through a series of convolution layers with Kernel's, Pooling layers, fully connected layers and then apply activation functions like SoftMax for classification of objects with probabilistic values between 0 and

1. The figure below depicts complete flow of CNN process, how an input image is taken and classification of objects is done.

An image is considered as a matrix of pixel values. By applying relevant filters, a Convolutional Neural Network can capture the Spatial and Temporal dependencies in an image successfully. Due to reduction of parameters involved and reusability of weights this architecture's performance is better on the image dataset. In simple words, the model can be trained to understand the image better.

The RGB images are generally separated by its three-color planes: Red, Green, and Blue. Images

ISSN: 2233-7857 IJFGCN

exist in many color spaces. Some of them are: Grayscale, Color (RGB), HSV, CMYK, etc.

Convolution Layer- Separable Convolution

The Kernel/Filter is the basic element involved in working of the convolution operation carried out in the Convolutional Layer.

ConvNet plays an important role in reducing the images into an easy form to process, so that the critical features for getting a good prediction are not lost. This is very important to design an architecture which is good at learning features and is scalable to extensive datasets.

The main objective of Convolution Operation is, to extract the high-level features like edges, from the given input image. Figure 3 depicts the convolution operation on a MxNx3 image. The first ConvLayer is responsible for capturing the Low-Level features such as edges, color, gradient orientation, etc. With added layers, the architecture adapts to the High-Level features as well, giving us a network, which has the wholesome understanding of images in the dataset, like how we would [27].

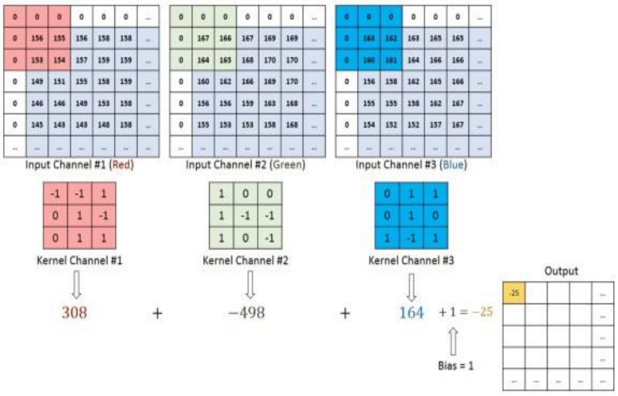
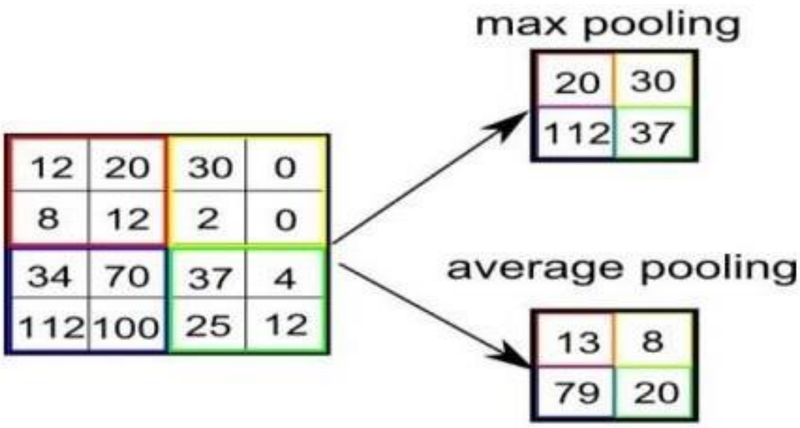


Figure 3. Convolution operation on a MxNx3 image matrix with a 3x3x3 Kerne

Pooling Layer

Like Convolutional Layer, to reduce the spatial size of the Convolved Feature, pooling layer is used. This is performed to reduce the computational power required to process the image through dimensionality reduction of the feature map. And it is used to extract dominant features from the image which are positional and rotational invariant, so that training of the model is done effectively.

Figure 4 describes the pooling types. Max Pooling and Average Pooling. Average Pooling returns the average of the values from each kernel covered by the portion of the image. On the other hand, Max Pooling extracts the maximum values from each of the kernel's covered by the portion of the image. It also acts as a Noise Suppressant. Whereas, Average Pooling only performs dimensionality reduction. Hence, the performance of Max Pooling is much better than Average Pooling.



ISSN: 2233-7857

Figure 4. Pooling Types

Classification : Fully Connected Layer

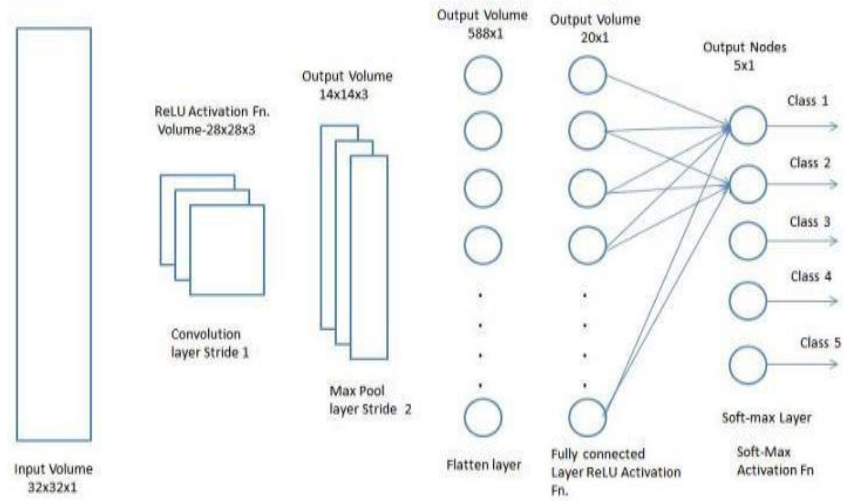


Figure 5. Fully connected layer

Now, as the image has been converted into an appropriate form for Multi-Level Perceptron, flattening the resulted image into column vector must be performed. The flattened output is then fed to a feed-forward Convolutional neural network and then backpropagation is applied to every iteration during training. After a series of epochs, the model can distinguish between dominating features and certain low-level features from the given input images and the classification is done using the SoftMax Classification technique. This is depicted in figure 5 shown above.

IV. RESULTS AND ANALYSIS

A model was developed and implemented which can identify the presence of Autism in children through their facial images by recognizing the facial landmarks and then classifying the behavior of the child through given image.

The Training and Testing is done on the Autistic children dataset considering two different variants of the data, providing the model 90% training data, 10% testing data and providing the model 80% training data, 20% testing data.

Dataset

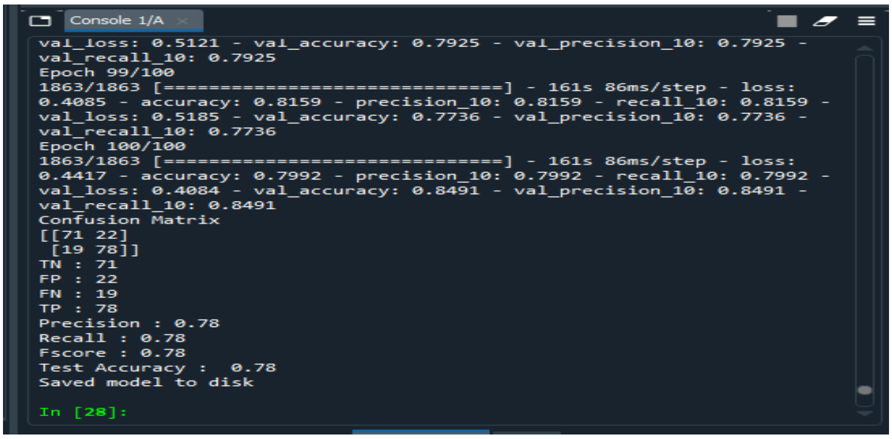
Early detection of Autism is extremely important to the child's development. This dataset was developed by Gerry Piosenka. There are many medical studies that are conducted which show a clear correlation between facial features and Autism. He tried to develop an accurate means by which parents that suspect their child might be autistic could submit one or better yet several images of their child's face to an internet website and get a probability estimate with respect to the presence of Autism. Autistic children Dataset was used from Kaggle.

- The data set consists of 4 directories.
- The train directory consists of 2 subdirectories, Autistic and Non_Autistic, each containing 1327 image files in jpg format. These are the files used for training the network
- The test directory contains 2 subdirectories labelled as above each containing 270 image files used to evaluate the model accuracy.
- The valid directory is similarly structured and contains 40 files in each subdirectory.

The fourth directory labelled consolidated consists of two subdirectories, Autistic and Non_Autistic. Each contains 1507 image files. It was created by merging the train, test and valid files into a single concatenated file

Results

Figure 6 shows the results of the CNN model with 90% Training data and 10% Testing data in terms of the performance measures of the CNN model. Test Accuracy, Precision, Recall, and Fscore.

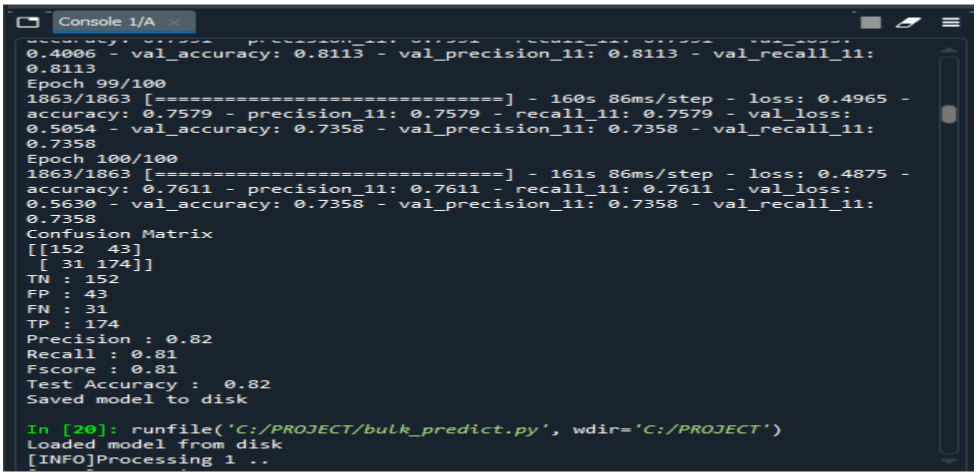


```
val_loss: 0.5121 - val_accuracy: 0.7925 - val_precision_10: 0.7925 -
val_recall_10: 0.7925
Epoch 99/100
1863/1863 [==============================] - 161s 86ms/step - loss:
0.4085 - accuracy: 0.8159 - precision_10: 0.8159 - recall_10: 0.8159 -
val_loss: 0.5185 - val_accuracy: 0.7736 - val_precision_10: 0.7736 -
val_recall_10: 0.7736
Epoch 100/100
1863/1863 [==============================] - 161s 86ms/step - loss:
0.4417 - accuracy: 0.7992 - precision_10: 0.7992 - recall_10: 0.7992 -
val_loss: 0.4084 - val_accuracy: 0.8491 - val_precision_10: 0.8491 -
val_recall_10: 0.8491
Confusion Matrix
[[71 22]
 [19 78]]
TN : 71
FP : 22
FN : 19
TP : 78
Precision : 0.78
Recall : 0.78
Fscore : 0.78
Test Accuracy :  0.78
Saved model to disk

In [28]:
```

0Figure 6. Output of Training Model with 90%-10% training and testing data

Figure 7 shows the results of the CNN model with 80% Training data and 20% Testing data in terms of the performance measures of the CNN model. Test Accuracy, Precision, Recall, and Fscore.



```
0.4006 - val_accuracy: 0.8113 - val_precision_11: 0.8113 - val_recall_11:
0.8113
Epoch 99/100
1863/1863 [==============================] - 160s 86ms/step - loss: 0.4965 -
accuracy: 0.7579 - precision_11: 0.7579 - recall_11: 0.7579 - val_loss:
0.5054 - val_accuracy: 0.7358 - val_precision_11: 0.7358 - val_recall_11:
0.7358
Epoch 100/100
1863/1863 [==============================] - 161s 86ms/step - loss: 0.4875 -
accuracy: 0.7611 - precision_11: 0.7611 - recall_11: 0.7611 - val_loss:
0.5630 - val_accuracy: 0.7358 - val_precision_11: 0.7358 - val_recall_11:
0.7358
Confusion Matrix
[[152  43]
 [ 31 174]]
TN : 152
FP : 43
FN : 31
TP : 174
Precision : 0.82
Recall : 0.81
Fscore : 0.81
Test Accuracy :  0.82
Saved model to disk

In [20]: runfile('C:/PROJECT/bulk_predict.py', wdir='C:/PROJECT')
Loaded model from disk
[INFO]Processing 1 ..
```

Figure 7. Output of Training model with 80%-20% training and testing data.

Below figures 8 and 9 shows the outputs of images with facial landmarks and the behavior of the child.

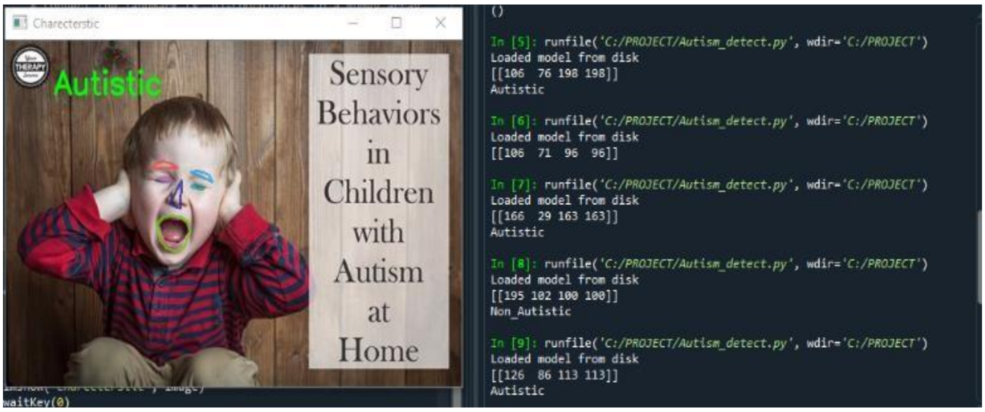


Figure 8. Testing the model by giving an input image.

Figure 9. Testing the model by giving an input image.

And, the testing is also performed on a set of images in a folder which is given as input for classification. These images are classified and stored in the result folder. Below figure depicts these results.

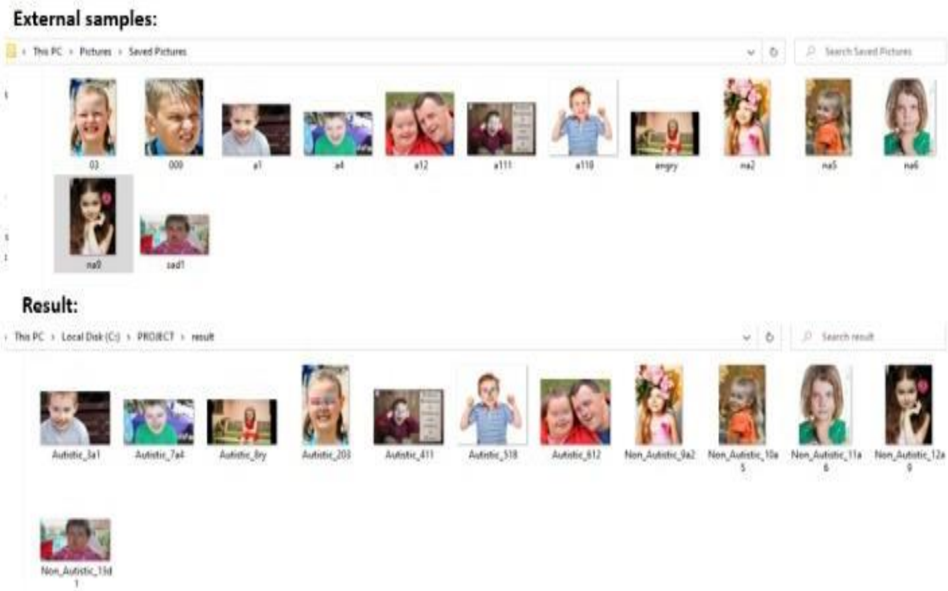


Figure 10. Test results of Internal and External Validation.

Confusion Matrix

The confusion matrix, in machine learning, is a grid of values that help to evaluate the performance of supervised classification models, depicted in Table 1. From this grid, you can also compute several metrics to give a score for the model these include precision, recall and the F1-score.

Actual			
Predicted	Total no: of Samples	Autistic	Non_Autistic
	Autistic	TP=True 0Positive	FP=False Positive
	Non_Autistic	FN=False Negative	TN=True Negative

Table 1. Confusion Matrix

Internal Validation and External Validation

Internal and external validation are concepts that reflect whether the results of a study are trustworthy and meaningful.

Internal Validation relates to how well a study is conducted. It is the study of a trustworthy cause-and-effect relationship established between training the model and its outcome. Internal validity depends largely on the procedures of a study and how rigorously it is performed.

External Validation relates to how applicable the findings are to the real world. It refers to how well the outcome of a study can be applied to other settings. That is, this type of validation refers to how generalizable the findings are.

In our model we define the validation in terms of Performance measures like Accuracy, Precision, Recall and Fscore and these metrics are based on the confusion matrix obtained as a result of training the model.

Accuracy

It may be defined as the number of correct predictions made by our ML model. We can easily calculate it by confusion matrix with the help of following formula:

Accuracy=TP+TN/(TP+FP+FN+TN).

Precision

Precision, used in document retrievals, may be defined as the number of correct documents returned by our ML model. We can easily calculate it by confusion matrix with the help of following formula:

Precision=TP/(TP+FP).

Recall

Recall may be defined as the number of positives returned by our ML model. We can easily calculate it by confusion matrix with the help of following formula:

Recall=TP/(TP+FN).

Below are the findings of the Internal validation and External validation of the implemented Model given in a tabular format.

INTERNAL VALIDATION			EXTERNAL VALIDATION		
N=400	Autistic	Non_Autistic	N=150	Autistic	Non_Autistic

<table>
<tr>
<td>
<table>
<tr><td>Autistic</td><td>TP=174</td><td>FP=43</td></tr>
<tr><td>Non_Autistic</td><td>FN=31</td><td>TN=152</td></tr>
</table>
Classification Rate/Accuracy
Accuracy = (TP + TN) / (TP + TN + FP + FN) = 0.82
Recall = TP / (TP + FN) = 0.81
Precision = TP / (TP + FP) =0.81
F-measure:
F-measure = (2 * Recall * Precision) / (Recall + Precision) = 0.81
</td>
<td>
<table>
<tr><td>Autistic</td><td>TP=63</td><td>FP=28</td></tr>
<tr><td>Non_Autistic</td><td>FN=17</td><td>TN=42</td></tr>
</table>
Classification Rate/Accuracy
Accuracy = (TP + TN) / (TP + TN + FP + FN) = 0.70
Recall = TP / (TP + FN) = 0.79
Precision = TP / (TP + FP) = 0.69
F-measure:
F-measure = (2 * Recall * Precision) / (Recall + Precision) = 0.73
</td>
</tr>
</table>

Table 2. Analysis of Internal and External Validation

A CNN Model is built to identify the presence of Autism in facial images. To validate the model, both internal and external validations are performed. The Results of the performance metrics in terms of Precision, Recall and Accuracy are projected in the above table 2.

The results exhibit the performance of our implementation is around 82% in terms of internal validation. A new dataset was created by downloading images from google and using this data, external validation is performed, and it is observed that the performance was around 70%.

V. CONCLUSION

ASD (Autism Spectrum Disorder) has its roots in very early brain development. However, the most understandable signs of autism and indicators of autism incline to begin between 2 and 3 years of age(Adams , 2011) Children with autism have often been reported to have gastrointestinal problems that are more frequent and severe than in children from the general population.

Children with this disorder may not have a good relationship with playing games. They tend to change the orientation of the game designed into their own. An example of this would be, lining up toys instead of playing with them. They also have no interest in participating in a "pretend game"; which plays a character in a group. This is a normal game among kids above age 2. One of the methods to detect autism in children is to analyze the interaction of a preschool kid with his/her toys. Many children with autism have long period of attention to a self-initiated activity. This is not true in the case of joint challenging activity with another person. Their activities are usually individual oriented. With the addition of their physical reaction to a stressful situation, such as hand flapping, prevents them from having social interactions with other people [10].

In conclusion, autism is a neuropsychological syndrome that confines the ability of individuals to socially interact with people and restrict their communication ability. But the early identification of this behavior plays a vibrant role in treating the children with ASD successfully.

Taken together, Findings show that, the presence of autism in children can be identified through Facial Images with an accuracy of 82%, and the measures of performance of the implemented model are calculated using Precision, Recall, F-score. And finally, the accuracy of the model is validated against internal and external validation as well.

VI. REFERENCES

[1] Ackovska, N., Kirandziska, V., Tanevska, A., Bozinovska, L., & Bozinovski, A. (2017). Robot - Assisted therapy for autistic children. SoutheastCon 2017. doi:10.1109/secon.2017.7925401.

[2] Adams, J. B., Johansen, L. J., Powell, L. D., Quig, D., & Rubin, R. A. (2011). Gastrointestinal

flora and gastrointestinal status in children with autism. BMC Gastroenterology,11(1). doi:10.1186/1471-230x-11-22.

[3] Altay, O., & Ulas, M. (2018). Prediction of the ASD Diagnosis with Linear Discriminant Analysis Classifier and K-Nearest Neighbor in Children. 2018 6th International Symposium on Digital Forensic and Security (ISDFS). Antalya, Turkey: IEEE. doi:10.1109/ISDFS.2018.8355354.

[4] Beibin Li, Sachin Mehta, Deepali Aneja, Claire Foster, Pamela Ventola, Frederick Shic, Linda Shapiro,"A Facial Affect Analysis System for Autism Spectrum Disorder" arXiv.org, cs, arXiv:1904.03616

[5] Eman Ahmed Zaky, "Autism Spectrum Disorder (ASD); The Past, The Present, and The Future", Child Adolesc Behav 2017, ISSN:2375-4494

[6] Frith, U. (n.d.). Asperger and his syndrome. Autism and Asperger Syndrome,1-36. doi:10.1017/cbo9780511526770.001

[7] Joseph Manfredonia, "Automatic Recognition of Posed Facial Expression of Emotion in Individuals with Autism Spectrum Disorder" Springer 2018, doi.org/10.1007/s10803-018-3757-9.

[8] Loth E, Spooren W, Ham LM, Isaac MB, Auriche- Benichou C, Banaschewski T, Baron-Cohen S, Broich K, Bolte S, Bourgeron T, et al. Identification and validation of biomarkers for autism spectrum disorders. Nat Rev Drug Discov. 2016;15:70–3.

[9] Lord, C., Risi, A. (2000). The Autism diagnostic observation schedule-Generic: A standard measure of social and communication deficits associated with the ASD. Journal of Autism and Developmental Disorders, 205-223.

[10] Mary Randolph-Gips, Autism: A Systems Biology Disease (2011). IEEE International Conference on Healthcare Informatics. DOI: 10.1109/HISB.2011.13

[11] Suzan Anwar, Mariofanna Milanova, "Real Time Face Expression Recognition of Children with Autism", ResearchGate, November 2016, IAEMR.

[12] Anthony P Atkinson, Winand H Dittrich, Andrew J Gemmell, Andrew W Young "Emotion perception from dynamic and static body expressions in point-light and full- light displays" Perception, volume 33, 2004, pp. 717- 746

[13] E. Fombonne, "Epidemiological surveys of autism and other pervasive developmental disorders: an update," J Autism Dev Disord, vol. 33, pp. 365-82, Aug 2003.

[14] M. D. Kogan, et al., "Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007," Pediatrics, vol. 124, pp. 1395-403, Nov 2009.

[15] A. P. Association, " Diagnostic and statistical manual of mental disorders (4th ed.)," A. P. Association, Ed., ed. Washington D.C., 1994.

[16] S. E. Folstein and B. Rosen-Sheidley, "Genetics of autism: complex aetiology for a heterogeneous disorder," Nat Rev Genet, vol. 2, pp. 943-55, Dec 2001.

[17] Frith, U. (n.d.). Asperger and his syndrome. Autism and Asperger Syndrome,1-36. doi:10.1017/cbo9780511526770.001

[18] Kanner, L. (1992). Follow-up Study of Eleven Autistic Children Originally Reported in 1943. Focus on Autistic Behavior, 7(5), 1-11. doi:10.1177/108835769200700501.

[19] O'Reilly B, Wicks K (2016) The complete autism handbook: The essential resource guide for autism spectrum disorder in Australia and New Zealand. Ventura Press, Sydney.

[20]http://www.parents.com/health/autism/history-of- autism/

[21]http://www.news-medical.net/health/Autism- History.aspx#

[22] http://www.bestbehaviour.ca/briefhistory.htm

[23]http://autismedsp5310s20f10.pbworks.com/f/The+histo ry+of+autism.pdf

[24]http://www.paulvangent.com/2016/08/05/emotion- recognition-using-facial-landmarks/

[25]https://towardsdatascience.com/a-comprehensive- guide-to-convolutional-neural-networks-the-eli5-way- 3bd2b1164a53 [26]https://www.pyimagesearch.com/2017/04/03/facial- landmarks-dlib-opencv-python/

[27]https://www.pyimagesearch.com/2018/02/26/face- detection-with-opencv-and-deep-learning/

[28] https://www.learnopencv.com/image-recognition-and- object-detection-part1/

[29]https://www.pyimagesearch.com/2014/11/10/histogra m-oriented-gradients-object-detection/

[30]https://towardsdatascience.com/a-guide-to-face- detection-in-python-3eab0f6b9fc1