# Modified Densenet for Face based Autism classification

**Vijayalaxmi Angadi and Dr Amuthabala**

*School of CSE, REVA University,* Bangalore, India
a.vijayalaxmi@gmail.com, amuthabala@gmail.com

## Abstract.

Computer aided diagnosis has become important in medical applications to aid the medical practitioner in making accurate decisions. Autism disorder classification from face image is one such application. Many machine learning algorithms have been proposed to classify Autism disorder with various levels of accuracy. In this work, a modified dense net neural network model is proposed to classify face image images to two classes of Autism disorder or no disorder. The original dense net model is extended with some connectivity changes to improve the accuracy of the model in this work. The modified dense net model is able to learn intricate features from the face image and able to classify it with higher accuracy compared to other deep learning models.

**Keywords**. Autism, Densenet, Face Image, Computer Aided Diagnosis

## 1. INTRODUCTION

Autism is a hereditary neurological disorder with complicated causes and courses. Person with Autism disorder has impaired social interaction and interpersonal communication. Individuals with Autism disorder have reduced ability in recognition of emotions. They have reduced attention to faces and diminished eye contact. As Face is an indicator of interpersonal interaction and communication, using Facial features to detect Autism disorder has gained importance. This work explores the use of computer aided diagnosis to detect autism from face images. Computer aided diagnosis (CAD) is a process to augment medical professional or in some case replace them in clinical diagnosis. They use various machine learning and image processing algorithms to diagnosis diseases. Typical CAD has stages as shown in Figure 1.

Acquisition is the first step and it is all about acquiring the image from sources using devices like camera, scanners, sensors etc. The devices provide raw data which must be processed using techniques like sampling and quantization to convert to a digital form for storage. The image acquired from previous step must be pre-processed and made suitable for further processing. The performance of further steps is dependent on the efficiency of image pre-processing. The type of pre-processing to be applied depends on the application requirements. Also depending on image is color or gray scale, the type of pre-processing to be applied to the image also varies. After pre-processing, the next stage is segmentation. . In this stage, homogenous parts in the image are identified and segregated. This segregation is done to identify similar areas of an image. Segmentation also segregates region of interest in the image. Segmentation is a important operation in various applications like computer vision, disease diagnosis etc. The segmented image is then represented in suitable form for further processing. Features like texture, intensity etc are extracted from the segmented objects. Machine learning based classifiers are then trained with the features to recognize the disease.

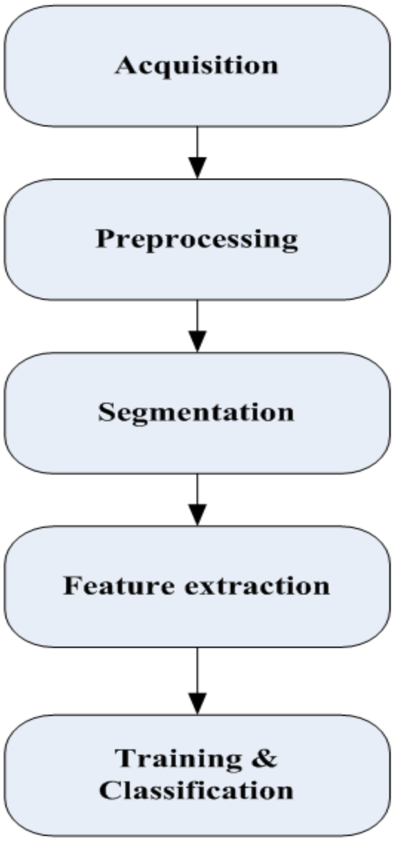


Fig. 1. Stages in Computer aided diagnosis

Machine learning is a algorithmic procedure which learns the patterns to recognize the classes in training stage and uses these patterns to classify any new cases. Deep learning is the last tend in machine learning where the features can be learnt automatically by the convolutional layers without manual description of features. By this way, deep learning models are able to learn intricate features from the images.

In this work, a deep learning model called dense net is adapted to classify the Autism disorder from face image. The original Densenet has more connections. The connections can lead to over learning and affect the accuracy. This work proposes connection optimizations in original Densenet for improved accuracy and reduction in classification time compared to original Densenet. The contributions of this work are as follows (i)A improved Densenet model for Autism disorder classification solving the problem of over learning in Densenet model

(ii)Demonstration of higher performance in proposed improved Densenet compared to state of art existing works.

## 2. RELATED WORK

Kumar et al [1] proposed a hybrid deep learning model for predicting children behavior based on emotions. Deep learning features extracted are classified using naïve Bayes and decision tree. The result is then fused to detect emotions. Sadiq et al [2] detected Autism using linguistic patterns uttered by people. The linguistic utterances of both disorder and normal patients are collected. MFCC features are extracted from the dataset and classified using LSTM deep learning model to Autism or normal condition. Maenner et al [3] extracted vocal features during word and phase utterances from children and classified it using Random forest classifier to recognize Autism. However false positives is higher in this approach. Liu et al [4] analyze the eye movement using machine learning to classify the children with/without Autism. Since the analysis is conducted in controlled environment with limited challenges, the method has higher false positives. Omar et al [5] proposed a Autism prediction model for any age by combining Random Forest CART and Random Forest ID3. The method was tested on AQ-10 dataset. Bi et al [6] extracted features from functional MRI and trained a multi class SVM to classify non AUTISM and various AUTISM types. The method can also classifies the abnormal brain regions. But data acquisition procedure is complex in this work. Jamal et al [7] proposed a machine learning based classification to recognize Autism from EEG samples. Phase synchronized patterns are extracted from EEG samples for three cases of fearful, happy and neutral for both autism and normal cases. The difference between minimal and maximal occurring synchrostates is used as brain connectivity features and SVM classified is trained with these features to classify Autism. The method works only for EEG collected in a controlled environment. Kong et al [8] extracted brain connectivity features from the T1-weighted MRI images and used deep neural network (DNN) classifier to recognize Autism. Though this method was able to detect Autism with accuracy of 90%, the data acquisition must be done in a controlled environment. Caratte et al [9] made a visual representation of eye tracking patterns and used it for Autism detection. PCA features were extracted from visual eye tracking images and features were classified using SVM classifier. Wu et al [10] detected Autism from videos of directed gaze towards face or object of interest. Authors proposed two models: image based model and facial behavior features based model. Authors also proposed a feature selection process to identify the most significant statistical behavioral features to avoid class imbalance problem. Though the method achieved 80% accuracy, it could not select the most suitable frames for Autism detection. Akter et al [11] experimented with different machine learning classifiers like Decision tree, Regression classifiers, Neural network, SVM, KNN and Xgboost. Datasets from Kaggle were used for experimentation. The study found that Logistic regression has higher accuracy compared to other classifier. Fawaz et al [12] experimented with various deep learning model for Autism detection from face images. Face images collected from Kaggle were classified using CNN, VGG16 and Xception. The method was able to achieve an accuracy of 91%. Yoleu et al [13] proposed a deep learning approach to recognize facial expressions. Facial expression detection is an important component of detection of neurological disorders inclusing Autism. Towards it, authors extracted local part based features with holistic facial information for robust facial expression detection. Haque et al [14] addressed the problem of facial expression detection in any environment. Deep convolutional neural network model was used for Facial expression detection with preprocessing of images for four different lighting conditions. Rudovic et al [15] integrated cultural background in face based Autism detection approaches to increase the accuracy. Small amount of information about target children is fused through transfer learning to achieve higher accuracy of Autism detection.

Though Deep learning has been used extensively for Face based Autism detection, still there is a research gap in selection of a best model which can learn discriminative features for Autism detection. In this work an improved Densenet model is proposed to solve this problem.

## 3. METHODS

The stages of the proposed solution are given in Fig 2.

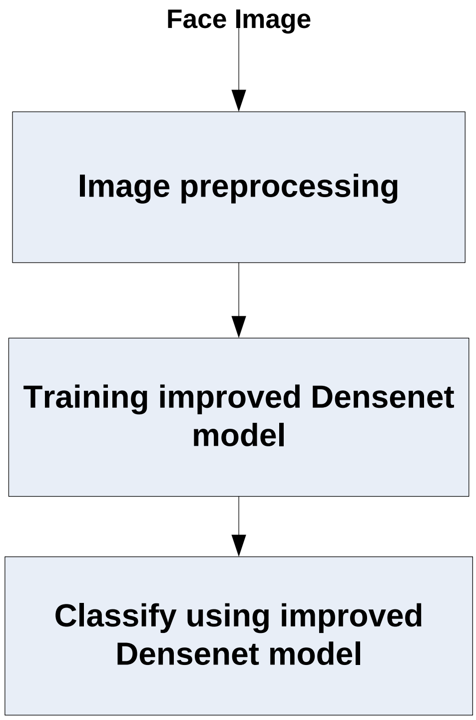


Fig 2. Stage of proposed solution

The Face images are first preprocessed during both training and classification before passing to Improved Densenet model.

Image is preprocessed by doing normalization, binarization, morphological operation and orientation. Normalization is a process that changes the range of pixel intensity values. The purpose of normalization is to bring image to range suitable for processing. Linear normalization is the process that changes the range of pixel values linearly, that is the input image is scaled linearly to scale required in the output image. In binarization process, the normalized image is converted to binary using OSTU thresholding. After binarization, morphological processing is applied on the image in terms of opening and closing operations. Opening is an image morphological operation that darkens small objects and entirely removes single-pixel objects like noise spikes and small spurs. The opening operation is erosion followed by dilation. Closing is an image morphological operation that completely removes noise involved within the object, and reduces the noise within the background. After morphological operation, the image is brought into standard orientation using random transform. After orientation, the image is resized to size of 224×224.

The preprocessed face images are taken to next stage of training the improved Densenet model.

Densenet is a recent deep learning model proposed for computer vision and object recognition applications. Similar to Resent, Densenet model is designed to solve the vanishing gradient problem. In addition, the convolution features of each layer are passed to subsequent layers as input in Densenet. Due to this, more intricate features can be learnt using the Densenet model.

The intricate features learnt are able to provide higher accuracy. The accuracy is generally lower in CNN with more layers due to vanishing gradient problem. In CNN with more layers, the features vanishes in the longer path travel and due to this, intricate features cannot be learnt.

A Densenet model has many dense blocks with varied number of filter. The dimensions are unique within the blocks. Between blocks, a transition layer is placed to do batch normalization. By this down sampling is done to match to dimensions of the subsequent layer. This work does changes in the Densenet model to improve the accuracy. The architecture of the improved Densenet model is given in Figure 2.

Fully connected layers are cut and replaced with fully convolutional layer. Also the pooling at layer 5 is removed to increase the stride by two times and compensates for edge localization. To attain hybrid features the convolutional layer is set to have a kernel of size $1 \times 1$ and channel depth 21. The resulting features from each convolution are accumulated with an additional layer. The feature map is up sampled at this additional layer with a convolutional layer of size $1 \times 1$. At up sampling layers, cross-entropy loss/sigmoid layer is attached.

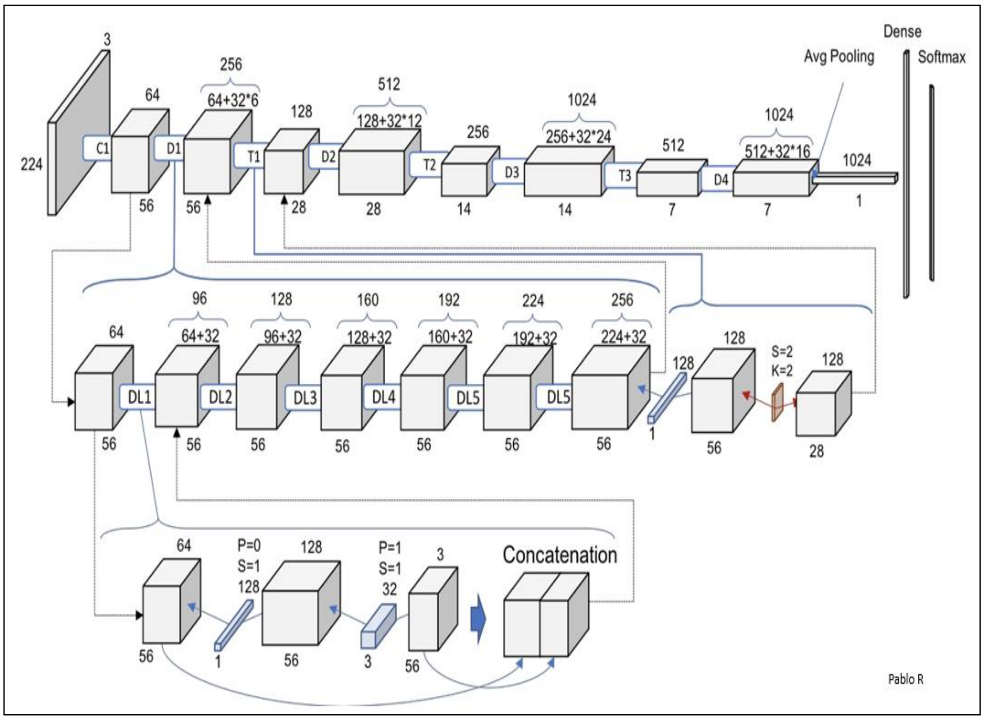


Fig. 3. Improved Denset model

A dataset of training images in two categories of Autism and normal is prepared. The dataset is split into training and test set in ratio of 80:20.. The training set images are used to train the improved Densenet model. The improved Densenet has following innovations.

(i).Fully connected layers are cut and replaced with fully convolutional layer. Also the pooling at layer 5 is removed to increase the stride by two times and compensates for edge localization

(ii) Updating the convolution operation with Local binary Pattern (LBP) feature based masking to improve the learning ability.

For a input gray scale image (q) of size 64×64 to the convolution operation, Local binary pattern (LBP) is computed. This LBP result and each of masks of default convolutions are joined with logical AND operation. Each of the 30 result after AND is then convolved with 7×7 kernel and summed up to get the output feature map. The output feature map with new convolution is given as

$$C(q)= \sum_{m=1}^{M} \sum_{n=1}^{N} [AND(LBP(q),mask(m) ).K] (j) \qquad (9)$$

Where M is the number of masks, N is the number of kernels and mask(m) is the binary mask of m^th pattern.

With this two innovations, the learning ability of the Densenet improves. It is able to identify more discriminative features and due to short circuiting, the classification time decreases.

After training, the test set images are passed to test the efficiency of the improved Densenet model.

## 4. RESULTS

The performance of proposed improved Densenet solution is tested using Autistic Children Facial image dataset collected from Kaggle[16]. The dataset has 1327 face images of Autistic children and 1327 face images of non Autistic children in the dataset. The dataset set is split to training and test set in ratio of 80:20.

The accuracy over various epochs in the proposed solution is given below

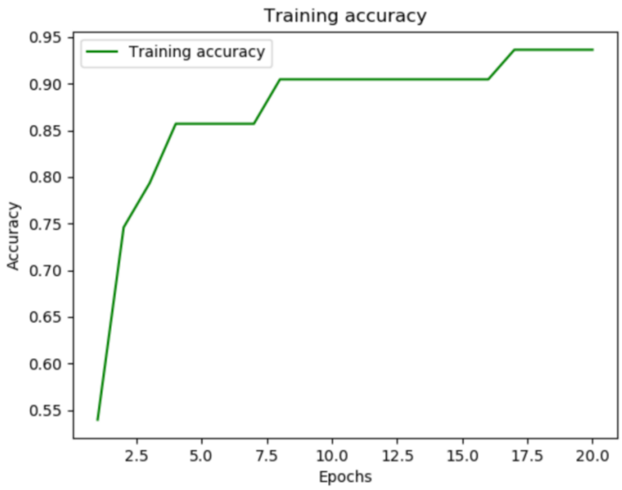


Fig. 4. Accuracy of proposed improved Densenet

The proposed solution was about to achieve about 93% accuracy at an epoch of 20 seconds. The training loss was measured for various epochs in proposed solution and the result is given below

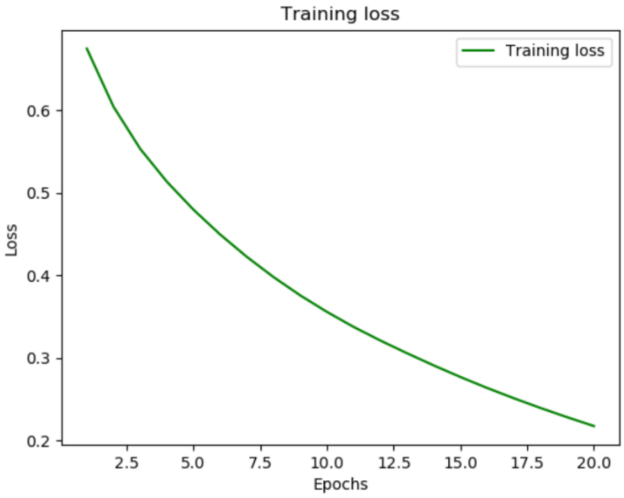


Fig. 5. Loss of proposed improved Densenet

The loss is minimal at 0.22 in epochs of 20. As the epoch increases, the loss decreases. It can be inferred that further increment in accuracy and reduction in loss can be achieved by increasing the training volume and increasing the epochs.

The performance of the proposed solution is compared against the deep learning model used in Fawaz et al : CNN, VGG16 and Xception

The comparison of accuracy and time for classification across recent works are given in Table I.

TABLE 1. COMPARISON OF PROPOSED SOLUTION

| **Model** | **Accuracy** | **Time (seconds)** |
|---|---|---|
| CNN | 0.85 | 0.21 |
| VGG16 | 0.87 | 0.23 |
| Xception | 0.89 | 0.37 |
| Proposed Densenet | 0.93 | 0.34 |

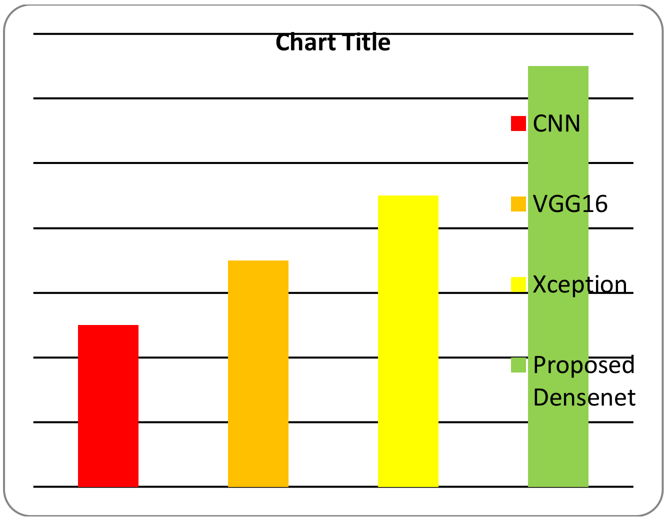


Fig. 6. Compariosn of accuracy

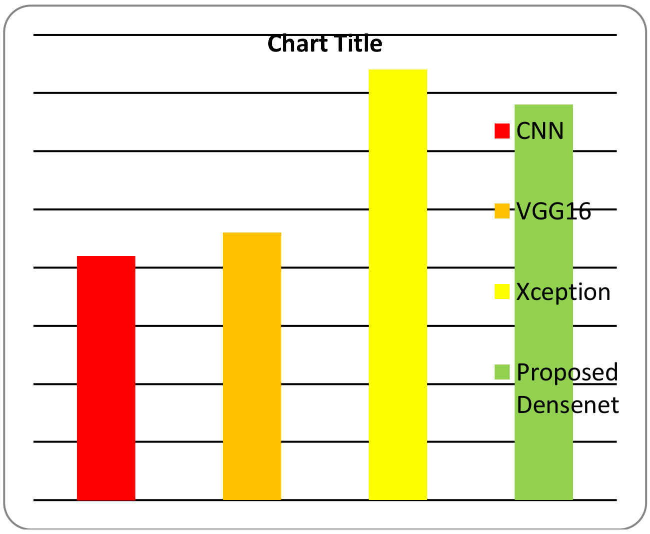


Fig. 7. Compariosn of classificaiton time

The proposed solution achieved an 8% higher accuracy compared to CNN, 6% higher compared to VGG16 and 4% higher compared to Xception. The accuracy has improved in the proposed solution due to two factors: better quality image due to preprocessing and more discriminative feature learning ability of improved Dense net. While other models like CNN, VGG16 and Xception are deep learning, the image was not enhanced before passing to deep learning stage and this affected the segmentation process.

The higher accuracy is achieved in proposed solution at cost of small increment in the time for classification compared to other models. As the number of interconnections is higher in the Densenet model compared to other deep learning models used in existing works like CNN, VGG16 and Xception, it has higher time compared to them, but due to optimization introduced, the proposed improved Densenet has lower time compared to original Densenet.

The comparison of performance of proposed solution with default Densenet (without any modification is compared) to gain insights into the effectiveness of proposed modifications and the result is given in Table 2

TABLE 2. COMPARISON WITH DENSENET

| **Model** | **Accuracy** | **Time (seconds)** |
|---|---|---|
| Default Densenet | 0.84 | 0.42 |
| Proposed Densenet | 0.93 | 0.34 |

The proposed improved Densenet is able to increase the accuracy by 9% and reduced execution time by 19% compared to default Densenet. This reduction is execution time is due to short circuiting certain interconnections to prevent from over fitting. Prevention of over fitting increased the accuracy by 9%.

The performance of the proposed Densenet model is tested with and without image preprocessing and the result is given below.

TABLE 3. IMAGE PREPROCESSING GAIN

| **Model** | **Accuracy** | **Time (seconds)** |
|---|---|---|
| Without Image processing | 0.88 | 0.39 |
| With Image preprocessing | 0.93 | 0.34 |

The proposed improved Densenet is able to achieve 5% more accuracy due to image preprocessing. The image preprocessing, is able to able to improve the contrast and localize the tumor patterns more effectively and this has increased the accuracy. The preprocessing operation has taken 12% higher computation time but considering the time is only 0.5 seconds and the accuracy gain of 5% this delay is tolerable.

## 5. CONCLUSION AND FUTURE WORK

An improved Densenet model is proposed to recognize Autism in this work. The proposed changes in Densenet model has helped to achieve a accuracy of 93% , which is 4% higher compared to existing deep learning models. Also the proposed changes have reduced the classification time by 0.03 seconds. Testing the effectiveness of proposed solution against larger dataset and extending the model for different age groups and culture is in scope of future work

## 6. REFERENCES


[1] A. Kumar and T. Senthil, "Construction of Hybrid Deep Learning Model for Predicting Children Behavior based on their Emotional Reaction", Journal of Information Technology, vol. 3, no. 01, pp. 29-43, 2021

[2] 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

[3] 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

[4] 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.

[5] 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.

[6] 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, pp. 18, Feb. 2018

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

[8] 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

[9] 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.

[10] C. Wu et al., "Machine Learning Based Autism Spectrum Disorder Detection from Videos," 2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM), 2021, pp. 1-6

[11] T. Akter, M. I. Khan, M. H. Ali, M. S. Satu, M. J. Uddin and M. A. Moni, "Improved Machine Learning based Classification Model for Early Autism Detection," 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), 2021, pp. 742-747

[12] Fawaz Waselallah Alsaade, Mohammed Saeed Alzahrani, "Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms ", Computational Intelligence and Neuroscience, vol. 2022

[13] G. Yolcu, I. Oztel, S. Kazan et al., “Deep learning-based facial expression recognition for monitoring neurological disorders,” in Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1652–1657, Kansas City, MO, USA, 13–16 November 2017.

[14] M. I. U. Haque and D. Valles, “A facial expression recognition approach using DCNN for autistic children to identify emotions,” in Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 546–551, Vancouver, Canada, 1–3 November 2018.

[15] Ahmed, S. T., Kumar, V. V., Singh, K. K., Singh, A., Muthukumaran, V., & Gupta, D. (2022). 6G enabled federated learning for secure IoMT resource recommendation and propagation analysis. *Computers and Electrical Engineering*, *102*, 108210.

[16] https://www.kaggle.com/general/123978

[17] E. Thirumagal, K. Saruladha,Chapter 2 - GAN models in natural language processing and image translation,Editor(s): Arun Solanki, Anand Nayyar, Mohd Naved,Generative Adversarial Networks for Image-to-Image Translation,Academic Press,2021,Pages 17-57,ISBN 9780128235195.

[18] T. E. and K. Saruladha, "Design of FCSE-GAN for Dissection of Brain Tumour in MRI," 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), 2020, pp. 1-6, doi: 10.1109/ICSTCEE49637.2020.9276797.