***https://doi.org/ 10.33472/AFJBS.6.5.2024. 3653-3662***

**African Journal of Biological Sciences**

# AUTISM SPECTRUM DISORDER PREDICTION IN CHILDREN USING ATTENTION BASED CONVOLUTION NEURAL NETWORK

***Mary Sebasti Rubala.S, Dr.Elamparithi.M and Dr.Anuratha.V,**

*Research Scholar, Department of Computer Science,
Kamalam College Of Arts and Science, Anthiyur,
Bharathiar University, Coimbatore, Tamil Nadu, India.
*E-mail: edupassrese.arch@gmail.com
Associate Professors, Department of Computer Science,
Kamalam College Of Arts and Science, Anthiyur,
Bharathiar University, Coimbatore, Tamil Nadu, India.

E-mail: profelamparithi@gmail.com and profanuratha@gmail.com

Article History
Volume 6, Issue 5, 2024
Received: 09 May 2024
Accepted: 17 May 2024
doi: *10.33472/AFJBS.6.5.2024. 3653-3662*

**Abstract:** Autism spectrum disease (ASD) is a neurodevelopmental disorder associated with abnormalities in brain development that impact facial appearance. Children with autism differ markedly from typically developed children in the patterns of their facial features. This work aims to provide a unique deep-learning method for diagnosing autism based on face cues utilizing a convolutional neural network with a multi-head attention (CNNMHA) mechanism, making diagnosing autism easier for families and clinicians. It mainly uses two phases such as preprocessing and classification. In this preprocessing step, the proposed system performs image resizing, image denoising by Gaussian filtering, and normalization to enhance the image quality and to make it more suitable for the specific classification task. In the classification stages, the proposed system uses CNNMHA to extract features from the preprocessed image and classifies the images into autistic children or non-autistic children. The data were obtained via Kaggle, and the preliminary computational results demonstrate that the proposed system outperformed existing methods. The technology can be used to help medical professionals validate their first screening results in order to identify youngsters with ASD illness.

***Keywords:*** *Autism spectrum disorder, Convolutional Neural Network, Gaussian Filtering, Deep Learning.*

## 1. INTRODUCTION

ASD is a neuro developmental disease that affects social communication and language skills throughout life. It is also characterized by limited and repetitive behavior [1]. Since ASD manifests itself differently in every person, it is referred to as a spectrum disorder. These deficiencies typically first appear in early childhood and result in difficulties functioning in various contexts [2]. To determine the association between autism and socio-demographic variables such as mother education, sex, age, and race, Leo Kanner initially classified autism in 1943. The World Health Organization reports that one child out of every 160 is diagnosed with autism globally each year. In wealthy nations, 1.5% of children received an ASD diagnosis in 2017 [3]. Various kinds of observations are used to identify the symptoms associated with ASD. When it comes to effective therapy for people with ASD, early detection does, however, require a substantial investment of time and energy [4]. To reduce the risk of ASD, a variety of therapies have been employed, including occupational therapy, behavioral analysis, physical therapy, speech therapy, and pharmaceutical therapy. However, the majority of these therapies have only had patchy success. Many (though not all) people with ASD need some form of lifetime assistance. Therefore, the need for innovative ASD therapies is essential [5].

In order to rapidly identify and assess ASD as well as other illnesses, including diabetes, stroke, and heart failure, numerous studies using different Machine Learning (ML) techniques have been carried out in recent years [6, 7]. Researchers use popular ML techniques like Support Vector Machines (SVM), Random Forests (RF), Naive Bayes (NB), K-nearest Neighbors (KNN), etc, to foresee meltdowns of autism [8]. They produce satisfactory outcomes for ASD prediction, but their performance and robustness are restricted by the size of the accompanying training data [9]. To mitigate these deficiencies, many researchers have recently used deep learning (DL) for ASD prediction. DL offers innovative healthcare applications. One of DL's main advantages is its ability to collect enormous volumes of data. To get the optimum result, it uses its sophisticated neural networks. Researchers and medical practitioners can use DL to uncover possibilities hidden in data [10]. This motivates us to propose an ASD prediction system using a deep learning approach with an attention mechanism. The manuscript's objectives are listed as follows:

- The study applies Gaussian filtering on the collected input images to suppress the noise, enhancing the classifier's prediction performance.

- The system uses the CNNMHA mechanism to extract the features and classifies the images as autistic or non-autistic. The Multi-Head Attention (MHA) mechanism, which emphasizes significant aspects and suppresses unimportant ones, can be considered a technique to identify the most informative, deeper features.

The remainder of the paperwork is sectioned as follows: The review of literature is presented in Section 2. Brief explanations of the suggested methodology are provided in Section 3. The outcome analysis of the methods is presented in Section 4. Finally, section 5 illustrates the articles' conclusion as well as future directions.

## 2. RELATED WORK

**Angelina Lu and Marek Perkowski [11]** recommended an ASD prediction system using DL. Firstly, the images were collected from the Kaggle Autism facial dataset and the East Asia ASD children's facial image dataset. After that, the system used a transfer learning model called visual geometry group-16 for prediction. The system attained an f-score and accuracy of 0.95 on the tested datasets. **Anupam Garg *et al.* [12]** suggested an explainable DL system for ASD prediction. To begin, the data was collected from the ASD screening dataset. The 15 features were randomly selected from the dataset, and finally, the ASD prediction was made using an explainable DL approach. The results showed that the system achieved 98% accuracy, which is better than the existing methods for the ASD screening dataset. **Kaushik Vakadkar *et al.* [13]** presented an ASD detection mechanism based on ML techniques. Initially, the preprocessing, such as missing values and outliers' removal, noise removal, encoding, and normalization, was performed on the collected dataset. After that, the system applied five classification models, namely LR, NB, SVM, KNN, and RF, on the preprocessed dataset for ASD prediction. The system used the publicly available Dr.Fadi Thabtahand dataset, and the LR achieved the highest accuracy of 97.15% compared to the other ML methods.

**Junxia Han *et al.* [14]** developed an ASD recognition system using a multimodal approach. First, a multimodal dataset was collected, and then the features like relative power energy, multi-scale entropy, brain network, and eye tracking were extracted. Finally, the stacked denoising auto-encoder algorithm was used for predicting ASD in children. Experimental results showed that the system achieved superior performance compared to the existing methods, i.e., achieved a maximum accuracy of 95.56%. **Irena Voinsky *et al.* [15]** proffered an RF model for ASD detection. Before classification, preprocessing was performed on the collected dataset to remove the null and missing values. The approach employed samples from 26 neurotypical controls and 73 ASD children from two cohorts (Israel and the USA) to predict ASD. According to the analysis, the algorithm produced results with an accuracy of 82% for correctly classifying children as either NT or ASD.

For the classification of autistic or non-autistic children, some of them use ML-based approaches. It offers efficient results when using SVM, LR, KNN, etc. However, these traditional techniques use hand-crafted features, which might not be the best feature set for categorization. These all have the potential to reduce the final performance. Therefore, to diagnose ASD more accurately, developing superior methods based on a larger dataset and employing an automated method to extract features is still required. The development of deep learning has allowed us to improve computer vision capabilities. It could use two-dimensional convolutional filters to extract task-specific characteristics automatically to get around the issue of preset features. So, this paper proposes a novel DL approach with an attention mechanism for ASD prediction that improves the system's performance.

## 3. PROPOSED METHODOLOGY

The workflow of the proposed methodology is displayed in Fig. 1. Two stages were mainly considered, namely, preprocessing and classification. To begin, the facial images are collected from the publicly available dataset. After that, the preprocessing, such as image resizing, image denoising, and normalization, is applied over the image, wherein the image denoising is carried out using the Gaussian filtering algorithm. Then, the classification of autistic or non-autistic children is done using the CNNMHA mechanism.

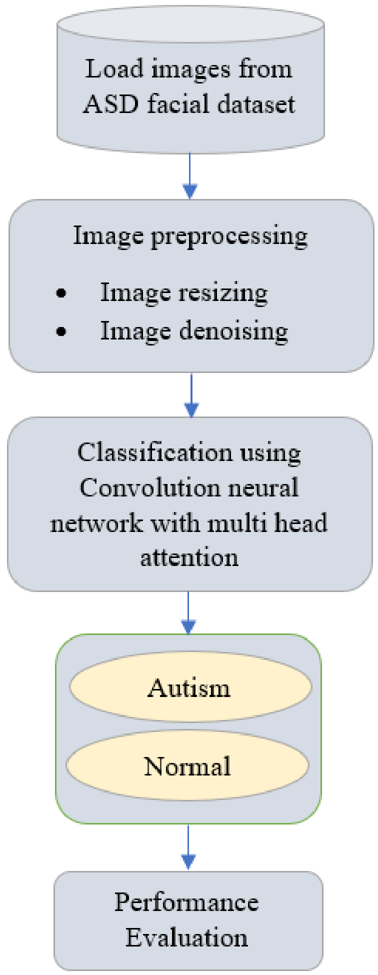


**Figure 1:** Workflow of the proposed methodology

### 3.1 Preprocessing

Initially, the system gathers data from a publicly available ASD facial dataset. After that, the dataset that was gathered was subjected to preprocessing. The gathered dataset contained some noise in facial images, which can lead to misclassification. Hence, the proposed system performs image resizing, image denoising, and normalization to improve the facial images' quality and make them more suitable for training and analysis. These are shortly described as follows:

**Step 1:** Image resizing

The majority of image collections contain shapes that are not suitable for feeding into the classifier. The suggested algorithms adhere to the size consistency requirement of 224 × 224 pixels or fewer for every input. Thus, the dataset was prepared for the proposed task by being transformed from random sizes to 224x224 pixels.

**Step 2:** Image denoising

The noise in the facial image is filtered using the Gaussian filtering algorithm. The Gaussian filter is one of the most commonly employed smoothing filters in image processing applications. The filtering function for a pixel in the image $(u, v)$ with the surrounding pixels' weighted average values is expressed as follows.

$$\overleftrightarrow{GF}(u, v) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2+v^2}{2\sigma^2}} \tag{1}$$

Where, $\sigma$ indicates the standard deviation of the distribution that is assumed to have a mean of 0.

**Step 3:** Normalization

Finally, the normalization is performed on the noise-removed images. The system rescales all imaging parameters from [0, 255] to [0, 1]. Normalization was done to prepare the dataset for Convolutional Neural Network (CNN) model training.

### 3.2 Classification

Next, the classification is done using CNNMHA, which classifies the preprocessed images into autistic and non-autistic based on facial features. CNN is typically shown as a series of fully connected and convolutional layers, with each layer's output subjected to a nonlinear activation function. The design may incorporate dropout layers and pooling to prevent overfitting. Robust features are automatically extracted by CNN and passed through multiple layers until the fully connected layer diagnoses ASD. CNN treated all of the preprocessed images of faces equally, without focusing on the images' deep features, even though each component of the images should have a different relevance. In order to address these issues, the suggested system uses MHA to learn the global feature relationship of input images adaptively. This reduces information redundancy among channels, learns the most crucial aspect of face images, and produces a more discriminative feature for ASD prediction. CNNMHA is the name used to describe this improvisation in traditional CNN. Thus, this improvisation in conventional CNN is termed CNNMHA. The structure of CNNMHA is shown in Figure 2. It consists of six layers: input, convolutional, activation, MHA, pooling, dropout, and fully connected. These are briefly explained as follows.

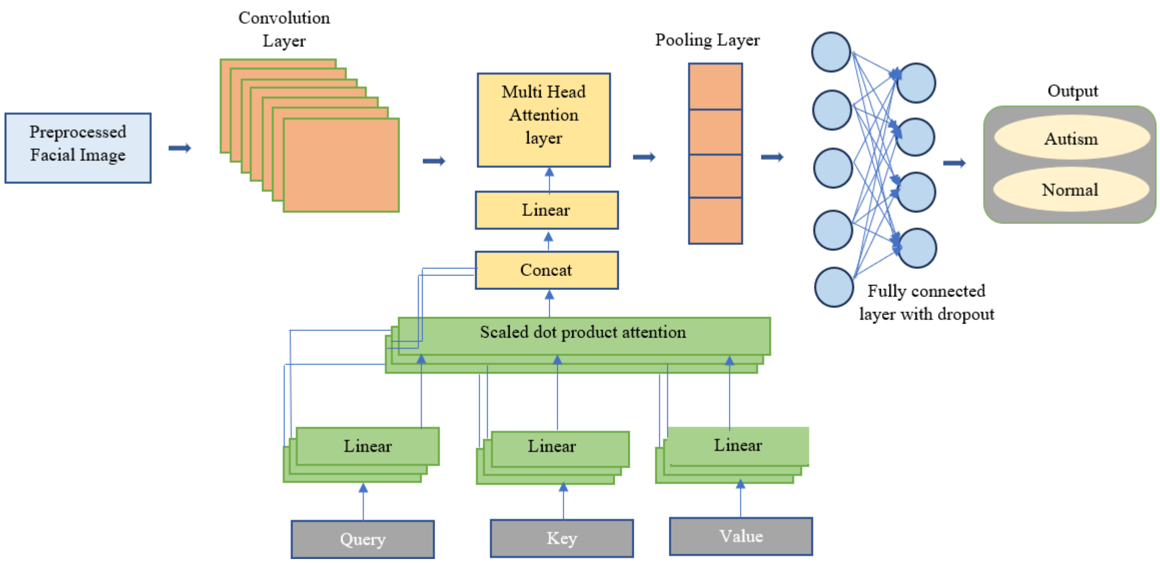


**Figure 2:** Structure of CNNMHA

## a) Input layer

It initially selects the preprocessed images as input for classification. It holds the image's pixel values.

## b) Convolutional layer

The pixel values of the image are passed to the convolutional layer. The convolutional layer is a crucial component of the CNN model utilized for feature extraction. The first convolutional layer learns essential characteristics like edges and lines. The features of squares and circles are extracted in the following layer. The face, eyes, and nose are the more intricate elements extracted in the subsequent layers. Convolutional layers use the same filters for every image pixel, lowering memory usage and improving efficiency.

The layer applies a group of learnable filters, also called kernels (mainly in the size of 2x2, 3x3, and 5x matrices) on the input data to obtain a feature map. After sliding across the input image data, the dot product of the kernel weight and the corresponding input image patch is computed. The volume that results from applying 12 filters to this layer will measure 32 × 32 x 12 feature maps.

## c) Activation layer

Activation layers introduce network non-linearity by including an activation function to the preceding layer's output. To address the gradient vanishing problems, an element-wise activation function is applied to the output of the convolution layer. ReLU is a frequently utilized CNN activation function. ReLU can overcome the gradient vanishing problems by maintaining the gradient without attenuation. It returns 0 if it receives any negative input; otherwise it returns the positive number $ip_g$. So the output of this layer will be 0 to infinity. It is mathematically expressed as follows:

$$\eta^* = \max\left(0, ip_g\right) \tag{2}$$

Where, $ip_g$ indicates the preprocessed image.

**d) MHA Layer**

Afterwards, the MHA layer receives the convolutional feature maps. MHA is a grouping of multiple heads that improves accuracy and performance of the system. It learns several deep distinguishing traits. Multiple heads are work with n-dimensional keys $(U_K)$, values $(U_V)$, queries $(U_Q)$. To obtain the output values, these keys, values, and queries are processed concurrently. The attention's output matrix for the input (convolutional feature maps) containing queries, keys, and dimension values $d_i$ is as follows:

$$Attention\left(U_Q, U_K, U_V\right) = soft\max\left(\frac{U_Q\, U_K{}^T}{\sqrt{d_i}}\right)U_V \tag{3}$$

Multiple "scalable dot product attention" is stacked to generate multi-head attention. By projecting the aforementioned queries, keys, and values linearly with $h_n$ times, it may thus see various sub-space representations at various points. Moreover, the aforementioned projection is carried out concurrently. The outputs can be written as

$$Multihead\left(U_Q, U_K, U_V\right) = concat\left(head_1, head_2, .....head_{h_n}\right)\widetilde{W}_m \tag{4}$$

Where, $\widetilde{W}_m$ refers to the weight matrices. The output from each head is concatenated to produce the final feature maps.

**e) Pooling Layer**

After that, the pooling layers will only down-sample the input along its spatial dimensions, resulting in activation with an even smaller number of parameters. The pooling layer creates a condensed feature map using each feature map's data from the MHA layer. The suggested method uses a 2 x 2 max pooling layer with strides two, and the output vector will be 16x16x12 in size.

**f) Dropout Layer**

A model's generalization capacity can be lost if it learns to memorize training material. During the network's training process, the dropout layer randomly removes nodes and connections to prevent over-learning. Every time the back-propagation algorithm modifies the network weight, the targeted dropout technique employs techniques to filter a collection of candidate weights. The random pruning procedure then applies this collection of candidate weights to the general dropout. Lastly, the model can pick up the skill of pruning throughout training to strengthen its resilience and increase its performance.

**g) Fully connected layer**

Eventually, fully connected layers are used for classification, following several convolutional, MHA, and pooling layers. The layer initially receives all the features and arranges them in a lengthy tube. The likelihood that input will belong to a specific class is, therefore, given by the softmax function of the layers. The softmax function generates 0 or 1 according to a binary classification. It is mathematically described as follows:

$$\hat{P}_R = Sigmoid\left(\overrightarrow{PF}\right) \tag{5}$$

$$\hat{O}_T^{''} = \begin{cases} 0, & \hat{P}_R \in (0, 0.5) \\ 1, & otherwise \end{cases} \tag{6}$$

Herein, $\hat{P}_R$ refers to the possibility that the review is a positive or negative, $\hat{O}_T^{''}$ indicates the final classification result where $\hat{O}_T^{''} = 0$ indicates the non-autism and $\hat{O}_T^{''} = 1$ indicates autistic children.

## 4. RESULTS AND DISCUSSION

Here, the simulation outcomes of the proposed system are analyzed with the existing systems regarding some performance measures. The Google Colab environment, which supports the most widely used deep learning libraries and machine setups with processing core 17 and 8GB RAM, was used to train the models in the cloud using Python. To test and validate the system, a publicly available autism facial image dataset from Kaggle is employed, and it is easily accessible through https://www.kaggle.com/datasets/cihan063/autism-image-data. There are 2940 facial images in this dataset: 1470 of the youngsters are autistic, and 1470 are not. Photographs of children with autism disorders were gathered from websites about the condition, while photographs of children without autism were selected at random from the Internet. From the total of 2940 images, 2540 images were taken for the training purpose, 300 images were for the testing process, and 100 images were for the validation purpose.

### 4.1 Performance Analysis

Here, the outcomes of the proposed CNNMHA are analyzed with the existing CNN, Artificial neural network (ANN), LR, and SVM concerning accuracy, precision, recall, f-measure, area under curve (AUC), false positive rate (FPR), false negative rate (FNR), and classification time.

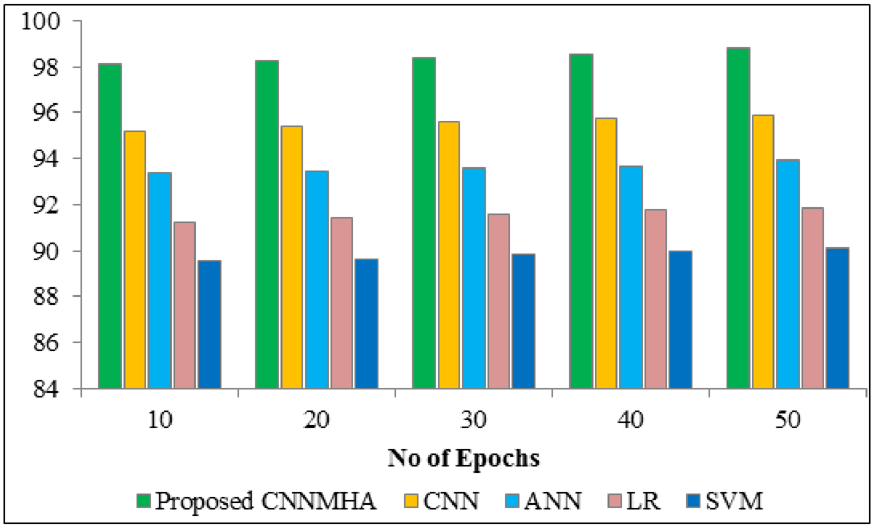


**Figure 2:** Accuracy analysis

Figure 2 demonstrates the proposed and existing accuracy for different epochs from 10 to 50. For ten epochs, the proposed model attained a maximum accuracy of 98.12%, whereas the existing models, such as CNN, ANN, LR, and SVM, yielded an accuracy of 95.23%, 93.36%, 91.23%, and 89.56%. Likewise, the proposed CNNMHA approach obtains the highest accuracy for all numbers of epochs (20 to 50) compared to the existing methods. Thus, it is confirmed that the proposed approach performs well for ASD classification in children compared to existing models. Table 1 demonstrates the outcomes of the methods concerning precision, recall, f-measure, FPR, FNR, and classification time.

**Table 1:** Results of the methods

| Metrics | Proposed | CNN | ANN | LR | SVM |
|---|---|---|---|---|---|
| Precision (%) | 98.52 | 95.67 | 93.72 | 91.69 | 89.92 |
| Recall (%) | 98.36 | 95.47 | 93.52 | 91.46 | 89.74 |
| F-Measure (%) | 98.44 | 95.47 | 93.52 | 91.46 | 89.74 |
| AUC (%) | 98.31 | 95.36 | 93.47 | 91.56 | 89.69 |
| FPR (%) | 0.021 | 0.085 | 0.108 | 0.642 | 0.903 |
| FNR (%) | 0.065 | 0.326 | 0.261 | 0.299 | 0.342 |
| classification time (min) | 0.85 | 1.01 | 1.95 | 2.28 | 2.99 |

The proposed method outperforms the traditional CNN, ANN, LR, and SVM approaches regarding all metrics. When considering the precision metric, the proposed CNNMHA attains 98.52% precision, whereas the existing CNN, ANN, LR, and SVM methods achieve the precision of

95.67%, 93.72%, 91.69%, and 89.92%, which are lower than CNNMHA. Likewise, the proposed method delivers better results for the remaining metrics. For example, the proposed one achieves 98.36% recall, 98.44% f-measure, 98.31% AUC, 0.021% FPR, and 0.065% FNR, better than the classical approaches. Also, the proposed one takes 0.85m to classify the autistic or non-autistic children, which is 2.14m less than the existing methods. The outcome analysis confirms that the proposed system is suitable for ASD detection in a minimal amount of time compared to the previous related techniques.

## 5. CONCLUSION

This study proposes an MHACNN for ASD prediction in children. The proposed system mainly comprises two parts: Image preprocessing and classification. The proposed system uses an ASD facial image dataset to test and verify the effectiveness of the system. The performance of the proposed CNNMHA is investigated against the classical approaches such as CNN, ANN, LR, and SVM approaches concerning accuracy, precision, recall, f-measure, AUC, FPR, FNR, and classification time. The results confirm the superiority of the proposed system over existing algorithms with higher accuracy and lower prediction time for ASD detection, i.e. the proposed MHACNN attains 98.43% maximum accuracy in 0.85 minutes. Thus, the overall experimental results showed the efficacy of the proposed work. In the future, we will use a transfer learning system to make a prediction even more accurate.

## COMPLIANCE WITH ETHICAL STANDARDS

This article does not contain any studies with human participants or animals performed by any of the authors.

## CONFLICT OF INTEREST

I, Mary Sebasti Rubala S declares no conflicts of Interest to disclose.

## REFERENCES

1. Sundas, A., Badotra, S., Rani, S., & Gyaang, R. (2023). Evaluation of autism spectrum disorder based on healthcare by using artificial intelligence strategies. *Journal of Sensors, 2023*, 1-12.
2. Feige, E., Mattingly, R., Pitts, T., & Smith, A. F. (2021). Autism spectrum disorder: Investigating predictive adaptive behavior skill deficits in young children. *Autism research and treatment, 2021*.
3. Alqaysi, M. E., Albahri, A. S., & Hamid, R. A. (2022). Diagnosis-based hybridization of multimedical tests and sociodemographic characteristics of autism spectrum disorder using artificial intelligence and machine learning techniques: a systematic review. *International Journal of Telemedicine and Applications, 2022*.
4. Bala, M., Ali, M. H., Satu, M. S., Hasan, K. F., & Moni, M. A. (2022). Efficient machine learning models for early-stage detection of autism spectrum disorder. *Algorithms*, *15*(5), 166.
5. Qiu, J., Kong, X., Li, J., Yang, J., Huang, Y., Huang, M., ... & Kong, J. (2021). Transcranial direct current stimulation (tDCS) over the left dorsolateral prefrontal cortex in children with autism spectrum disorder (ASD). *Neural plasticity, 2021*.

6. Hasan, S. M., Uddin, M. P., Al Mamun, M., Sharif, M. I., Ulhaq, A., & Krishnamoorthy, G. (2022). A Machine Learning Framework for Early-Stage Detection of Autism Spectrum Disorders. *IEEE Access*, *11*, 15038-15057.
7. Hanif, M. K., Ashraf, N., Sarwar, M. U., Adinew, D. M., & Yaqoob, R. (2022). Employing machine learning-based predictive analytical approaches to classify autism spectrum disorder types. *Complexity*, *2022*, 1-10.
8. Karim, S., Akter, N., Patwary, M. J., & Islam, M. R. (2021, November). A review on predicting autism spectrum disorder (asd) meltdown using machine learning algorithms. In *2021 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)* (pp. 1-6). IEEE.
9. Raja, K. C., &Kannimuthu, S. (2023). Conditional Generative Adversarial Network Approach for Autism Prediction. *Computer Systems Science & Engineering*, *44*(1).
10. Alam, S., Raja, P., & Gulzar, Y. (2022). Investigation of machine learning methods for early prediction of neurodevelopmental disorders in children. *Wireless Communications and Mobile Computing*, *2022*.
11. Lu, A., & Perkowski, M. (2021). Deep learning approach for screening autism spectrum disorder in children with facial images and analysis of ethnoracial factors in model development and application. *Brain Sciences*, *11*(11), 1446.
12. Garg, A., Parashar, A., Barman, D., Jain, S., Singhal, D., Masud, M., &Abouhawwash, M. (2022). Autism spectrum disorder prediction by an explainable deep learning approach. *Computers, Materials & Continua*, *71*(1), 1459-1471.
13. Vakadkar, K., Purkayastha, D., & Krishnan, D. (2021). Detection of autism spectrum disorder in children using machine learning techniques. *SN Computer Science*, *2*, 1-9.
14. Han, J., Jiang, G., Ouyang, G., & Li, X. (2022). A multimodal approach for identifying autism spectrum disorders in children. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *30*, 2003-2011.
15. Voinsky, I., Fridland, O. Y., Aran, A., Frye, R. E., & Gurwitz, D. (2023). Machine learning-based blood RNA signature for diagnosis of autism spectrum disorder. *International Journal of Molecular Sciences*, *24*(3), 2082.