Journal of Current Research in Engineering and Science
JCRES
Bi-Annual Online Journal (ISSN: 2581 - 611X)

VOLUME:7, ISSUE:2, PAPER NUMBER:44

Early Detection of Autism Spectrum Disorders Using Optimized Feature Selection and Deep Belief Network Generative Model

S. Kanimozhi[1], Dr T. Vimala[2], Dr Samundeshwari[3], Selva Jothi M[4]

[1]Department of Electronics and Communication Engineering, Chennai Institute of Technology,

Research Scholar, Dr.MGR Educational and Research Institute

[2]Department of Electronics and communication engineering, Dr.MGR Educational and Research Institute

[3]Department of computer science and engineering, Sri Sairam engineering College

[4]Department of Information Technology, Chennai Institute of Technology

Corresponding Author:

S. Kanimozhi;

Kanimozhi274@gmail.com

ABSTRACT

Patients with neurodevelopmental disorders such as autism spectrum disorder (ASD) experience difficulties in their daily lives. ASD affects children's behavior, learning abilities, social interaction, and communication. It is not curable. It impacts a number of behavioral domains, such as repetitive and stereotyped conduct as well as social and language ability. This disease is a neurodevelopmental ailment of great concern. The symptoms of ASD are remarkably similar to those of many other mental diseases, making the diagnosis of ASD challenging and time-consuming. With the increasing use of machine learning-based models in the prediction of numerous human diseases, early diagnosis based on various physiological and health variables are possible. The objective of this research is to develop a classification model with the highest level of accuracy possible for predicting the likelihood of ASD. We employed deep belief network generative models (DBNGM) and optimized feature selection dragonfly algorithm (DA) to explore the possibilities of predicting and assessing ASD features (ASDD-DADBNGM). These thorough experimental analyses suggest that choosing features wisely can be crucial to predicting ASD. We contend that the thorough feature importance analysis presented in this work will help medical professionals make decisions when they are screening patients for ASD. When compared to other methods currently used for the early identification of ASD, the suggested framework has produced encouraging outcomes.

Keywords:

Autism Spectrum Disorders, Disease Diagnosis, Deep learning, deep belief network, generative adversarial networks

I. INTRODUCTION

Autism spectrum disorder (ASD) is a type of neurological condition that mostly affects a child's ability to communicate and behave socially. It is identified as a significant mental illness in early childhood. For the rest of their lives, it continues. More than one percent of people worldwide suffer with autism [1]. In order to provide the best medical advice for improved patient health, significant research is required for the early diagnosis of ASD [2]. ASD mostly affects communication and social interactions, and it causes patients to engage in repetitive behaviors. Autism in youngsters is often discovered too late in developed nations. Autism is a disorder in which a patient's brain growth is inadequate, causing behavioral and social difficulties. These kids struggle with precise comprehension and learning, and their primary areas of difficulty are communication and socialization [3].

The condition usually appears before the child turns three, can be difficult to diagnose, and may persist for the remainder of the child's life [4]. Therefore, early identification of ASD can help with the previously listed problems and improve the quality of life for the individual with ASD and their family. Around the world, one in 160 children has ASD, according to data from the World Health Organization (WHO) [5]. Given the rising prevalence of ASD, early diagnosis is essential to mitigate its long-term implications.

In recent years, medical imaging analysis and illness detection have greatly improved thanks to machine learning and image processing algorithms, which now perform almost as well as medical professionals [6]. Concurrently, the employment of deep learning technologies in the healthcare business is an ongoing research topic [7]. The ability to access large volumes of data, freshly developed learning algorithms, and ever-increasing computer power make these advancements possible [7].

With the help of brain MRI data, this study attempts to develop an automated detection framework that can categorize ASD. The article proposes deep belief network generative models and optimized feature selection to explore the possibilities of predicting autistic patients and healthy individuals. Based on considerable experimental data, the present study outperforms the prior studies in terms of accuracy.

The contribution of the study;

- To improve the ASD grading accuracy, an effective deep learning algorithm has to be technologically improved using enhanced dragonfly optimization algorithm for irrelevant or repetitive features.
- To propose a neural-network based classification model to enhance the ASD detection at the earliest stage which having the highest probability of getting cured.
- To enhance the classification accuracy and reduce the error rate while classifying the ASD dataset.
- Evaluation is performed against traditional deep learning model with the proposed hybrid model with respect to accuracy and error metrics.

II. RELATED WORKS

Researchers have recently focused a lot of emphasis on deep learning-based classification algorithms because of their capacity to automatically identify features and make an excellent diagnosis of ASD [8].

In order to automatically diagnose ASD, Ma et al.'s [9] introduced a multi scale learning techniques to captures deep feature and constantly changing representations of rs-fMRI data. In order to gather the multi-scale statistical data, the author first splits the brain into multi-scale ROIs. We create multi-scale dynamic FCNs by employing sliding windows to partition fMRI time data in the temporal dimension, taking into account the dynamic nature of FCNs.

K. Devika [10] uses GAN encoder-decoder architecture to identify ASD using sMRI data. ASD diagnosis is made possible by the GAN, which was trained only on healthy scans and perceives ASD samples as anomalies. It is not necessary to use numerous ASD training samples with this strategy. The ABIDE II dataset is used to assess the model.

A CNN architecture for the automatic detection of ASD utilizing the CC400 functional parcellation and the ABIDE dataset was proposed by Sherkatghanad et al., [11]. Slice timing correction, motion correction, and voxel normalization were among the preprocessing processes used on MRI images that were run through 400 filters of varying sizes in CNN. The multilayer perceptron was given the entire CNN result to finish the classification process, resulting in an accuracy of 70.22%.

A deep learning (DL)-based system was suggested by Shin et al. [12] to identify children with ADHD who also have ASD. Thirteen ADHD children with co-occurring ASD and fifteen normally developing (TD) children recorded their functional near-infrared spectroscopy (fNIRs) signals while they drew handwriting patterns. By combining CNN and bidirectional long short-term memory (Bi-LSTM), a hybrid method was developed to identify children with ADHD who also had ASD. The experimental findings demonstrated that our hybrid method could detect

ADHD children with co-occurring ASD with 94.0% classification accuracy, 89.7% sensitivity, and 97.8% specificity based on the PL predict task.

Thermal imaging has been proposed by Rusli et al. [13] as a passive means of non-intrusively analyzing physiological data related to affective moods. According to the study's hypothesis, the various affective states experienced by autistic children may be directly impacted by changes in cutaneous temperature caused by pulsing blood flow in the blood vessels at the frontal face area as assessed by the modality. To quantify the thermal imaging data produced by various affective state expressions elicited by varying combinations of audio-visual stimuli, an organized experimental setting was created. The changes measured from the region of interest were identified using a wavelet-based method for pattern recognition in time series.

CNN with Residual Network was first introduced by Jain et al. [14] to classify autistic disorder using magnetic resonance imaging (MRI) brain pictures. To facilitate object localization and delineation, the cortical areas are segmented using hybrid FCM-GMM. The system offers enhanced contextual information by utilizing the probability distribution between the background areas and the brain image. VGG16 is used to extract features, improving classification accuracy and providing structural information. With the use of identity mapping and DCNN-ResNet based classification, the generalization problem is solved. An increased layer count makes training easier to complete and less error-prone.

III. METHODOLOGY

This study attempts to develop an efficient prediction model that detects autism in individuals by utilizing several machine learning techniques. This section explains the classifiers used in this research project and gives an overview of the ABIDE datasets. Then, evaluations are conducted on all performance factors related to early age and performance. For any kind of data, performance metrics and assessment standards are universally relevant. Presenting the suggested research methodology together with thorough justifications and illustrations brings this section to a close.

A. DATASET

The public dataset known as Autism Brain Image Data Exchange (ABIDE) is where the dataset was acquired [12]. Data on functional and structural brain imaging from 17 different international universities have been merged into a multisite platform known as ABIDE. The scientific community can access it, and each subject's anatomical and phenotypic data are provided together with resting state functional magnetic resonance imaging (rs-fMRI). There are 539 people with ASD and 573 typical controls in the database. This study combines normal controls with equal-sized rs-fMRI data from three different subtypes of autism: ATD, APD, and PDD. Every class has 36 subjects, each with its own student body.

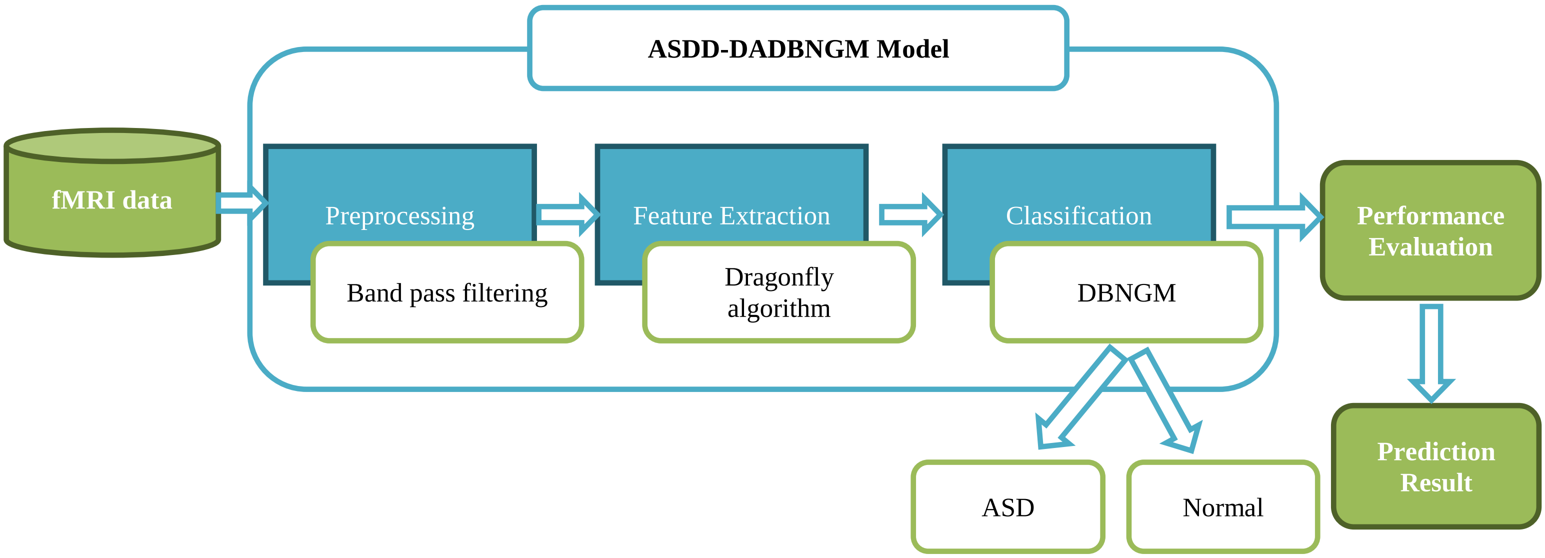


Figure 1. Proposed framework of DADBNGM Autism detection model

B. PREPROCESSING

The best way to preprocess the rs-fMRI data is still up for debate. To facilitate better data-sharing, the preprocessed version of ABIDE is made publicly available. Preprocessed rs-fMRI data (DPARSF) was obtained utilizing multiple pipelines from the Preprocessed Connectomes Project (PCP), which offered preprocessed versions of ABIDE-I [15]. The rs-fMRI data processing (MNI152) takes care of motion correction, spatial smoothing, normalization, and template space registration. To reduce the signal noise, band pass filtering (0.01–0.1 Hz), motion parameters and a global signal regression were employed.

C. DRAGONFLY ALGORITHM FEATURE SELECTION

The DA algorithm draws inspiration from dragonflies, as implied by its name [16].This approach can be viewed as a swarm intelligence method for figuring out an optimization problem's global optimum. The following lists the swarming behavior of dragonflies along with the mathematical models needed to simulate it [17]:

The strategy individuals use to keep themselves from colliding with their neighbors is known as separation. Equation (1) provides a mathematical representation of this phenomenon.

$$R_a = \sum_{b}^{n} P - P_b$$

where n is the neighborhood size, P_b is the b-th neighboring person of the P position, and P is the current individual's location. When two people are aligned, their velocities match those of other nearby people. Equation (2) provides a mathematical representation of this phenomenon.

$$X_a = \frac{\sum_{b=1}^{n} S_b}{n}$$

where n is the neighborhood size and S_b is the velocity of the b-th neighborhood individual. The term cohesion describes a person's propensity to be the center of mass in their neighborhood. The mathematical model for this behavior is found in Eq. (3)

$$X_a = \frac{\sum_{b=1}^{n} S_b}{n} - P$$

Other essential actions that every person engages in in order to survive are attraction to the food source and running away from enemies. Equation (4) models the attraction towards the food.

$$G_a = P^+ - P$$

where P^+ denotes the location of the food source and P is the position of the person at that moment.

$$H_a = P^- + P$$

where P denotes the position of the current person and P^- represents the position of the opponent.

DA used the step and position vectors as two vectors to solve optimization problems. It is defined which two vectors are. The definition of the step vector is as follows:

$$\Delta P_{t+1} = (rR_a + xX_a + cC_a + gG_a + hH_a) + w\Delta P_t$$

where the separation weight is denoted by r. R_a is the a-th individual's separation, x is the alignment weight, X_a represents the i-th individual's alignment, c denotes the cohesion weight, g is the food component, C_a is the cohesiveness of the a-th individual, The a-th individual's food supply is G_a, the adversary factor is h, and The iteration number is t, the inertia weight is w, and the opponent location of the a-th individual is represented by the symbol H_a.

$$P_{t+1} = \begin{cases} \neg P_{t,} \, p < T(\Delta p_{t+1}) \\ P_t, P \geq \ T(\Delta p_{t+1}) \end{cases}$$

where r is a random number in the range [0, 1], $T(\Delta p_{t+1})$ is calculated as in Eq. (8).

$$T(\Delta P) = \left| \frac{\Delta p}{\sqrt{\Delta p^2 + 1}} \right|$$

D. CLASSIFICATION MODEL- DBNGM

Deep belief network (DBN) is the most popular DL methods; it was first proposed by Hinton [18]. This algorithm finds the ideal settings more quickly than the others and is quick to learn new information. A logistic regression layer and an unsupervised learning module based on restricted Boltzmann machines (RBMs) are the fundamental components of a traditional DBN [19].

Layer-wise training of a well-known stochastic neural network, known as the RBM, creates a DBN. The RBM consists of two layers: one hidden layer of Boolean neurons and another layer of binary-valued neurons. The connections amongst neurons within a layer are not bidirectional or symmetrical, despite being between layers.

Layer-wise configurations use the energy function of the configuration, given in (2), to discover the probability distribution that exists between the two levels.his leaves us with the following equation to express the probability distribution.

$$Ef(a,b) = -\sum_{m=1}^{z_a} x_m a_m - \sum_{n=1}^{z_b} y_n b_n - \sum_{m=1}^{z_a} \sum_{n=1}^{z_b} b_n W_{n,m} b_m \tag{2}$$

$$pd = \frac{e^{-Ef(a,b)}}{\sum_a \sum_b e^{-Ef(a,b)}}$$

In the hidden layer, there are b_n Boolean hidden neurons, whereas $W_{n,m}$, neurons make up the visible layer. The weight matrices that divide the two layers are b_m and b_n. The biases for the two layers are $x_m$and y_n.

After that, an equation is created to express the activation probability functions.

$$pd(a_m = 1|b) = sig\left(\alpha_m + \sum_{n=1}^{i_b} W_{n,m} b_m\right)$$

$$pd(b_m = 1|a) = sig\left(y_m + \sum_{m=1}^{i_a} W_{n,m} a_n\right)$$

The logistic sigmoid function is also represented as sig(). The pre-training principles support this since the weight matrices and layer biases can be taught without supervision. The idiosyncrasies of the data are too complex for a single hidden RBM. Deep features from the input dataset can be gradually extracted by a DBN, which is built by stacking layers of RMBs in a hierarchical manner and ending with a logistic regression layer. The DBN's first RBM is pre-trained to function as an independent RBM using the training data as inputs.

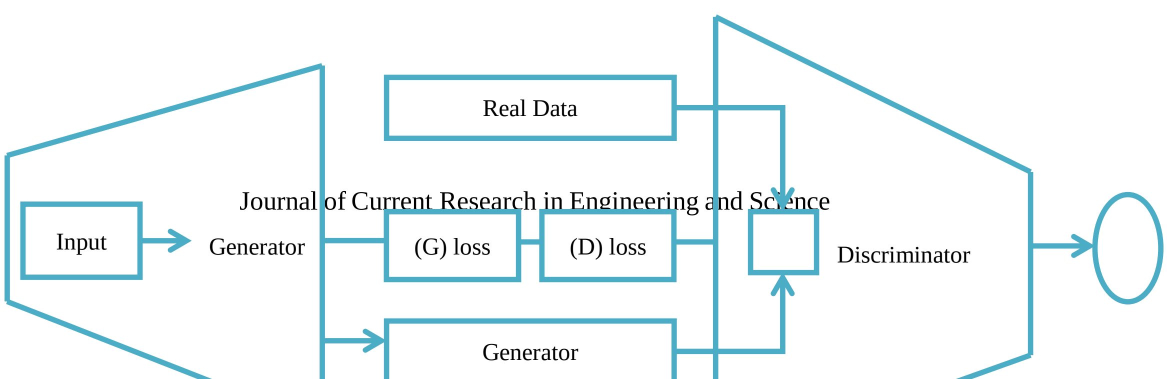
The output of the first RBM is chosen to be the input for the second RBM once the weight matrix and bias settings are determined. Then, using the same process, the invisible layers of the previous two RBMs are repeatedly trained to create a new RBM. The last stage involves overlaying a comprehensive predictor (such a logistic regression layer) over the network and closely monitoring its training. Following the application of the previously stated stages, the trained network's parameters are slightly altered by fine tuning using the back-propagation (BP) technique.

Using training sets that are similar to the training sets, GAN learns how to generate fresh data. The architecture of a GAN is shown in Figure 1. A discriminator (D) and a generator (G) are its two essential components. The discriminator is a model used to categorize instances as real (from the domain) or fake (manufactured), whereas the generator is a model used to generate fresh, persuasive examples from the problem domain.

A generator network is used to generate a sample of data instantly that looks like real data. Its competitor, the discriminator network, on the other hand, can differentiate samples obtained from the generator framework from samples obtained from real data. Classification models act as the discriminant. According to (Uddin, 2019), the objective function of the GAN design is in Eq. (1).

$$arg \min_{Gm} \max_{Dm} P(Dm, Gm) = E_{i\sim k_{data}(i)}\left[\log\left(Dm(i)\right)\right] + E_{a\sim k_a(a)}\left[\log\left(1 - Dm\left(Gm(a)\right)\right)\right]$$

where the discriminator function is denoted by $Dm(i)$. The result of this function is the probability that the input vector denoted as i comes from the training dataset. Given an input of i, the $Dm(i)$ function returns a number between 0 and 1.



Similar to this, $Gm(a)$ is a generator function that creates a matrix with the same dimensions as x depending on the z (noise vector). The training dataset is used to represent the distribution of sample chances, which is denoted by $k_{data}(i)$. The sample distribution of chances from a noise generator is denoted by the symbol $k_a(a)$.

The expectation function, often known as E(.), is a positive class that the log-loss function generates. The log loss function is defined in Equation (2).

$$E(s|\, r) = \frac{-1}{L}\sum_{x=1}^{L}(r_x(\log s_x) + (1 - r_x)(1 - s_x))$$

Whereas the real data is displayed as r_x, the estimation is represented as s_x. When it is expected that the response of the model would be either 0 or 1, the log function is utilized. When x is removed from (i), the probability distribution (i) indicates that (fi) of the provided function (i) is expressed as in Eq.

$$E_{i\sim s}(f(i)) \int s(i)f(i)dx$$

There are two loops in Eq. (1), namely $max_{DM}P(dm, gm)$ and $min_{GM}\ P(dm, gm)$. By adjusting discriminator parameters, $max_{DM}P(dm, gm)$ aims to maximize the right side. Two loops in the goal function shown by Equation (1) stand for $max_{DM}P(dm, gm)$ and $min_{GM}P(dm, gm)$. $min_{GM}P(dm, gm)$ seeks to minimize through parameter adjustments for the generator.

IV. EXPERIMENTAL SETUP

This section addresses the necessary hardware and software to run the model efficiently and is related to the result. The suggested work uses machine learning algorithms to enhance prediction quality on the previously given dataset. Evaluation measures, including as the accuracy, precision, recall, and F1-score, are used in the analysis of experiments to assess the model's performance and provide recommendations for the optimal strategy. Using the characteristics that were obtained from the dragonfly feature extractor, a data-driven approach is utilized in this work to identify ASD patients since the rs-fMRI data may indicate the functional connectivity links across the brain's areas[7]. This study contrasts the DBNGM categorization of ASD with the DBN and GAN classifications. We must employ a variety of datasets for both training and testing in order to appropriately evaluate the classifier's performance while generalizing to new data.

Table 1: Performance of Proposed DBNGM methods

Method	Accuracy	Precision	Recall	F1 score
DBN	96.74	96.12	96.89	97.23
GAN	98.76	98.54	98.73	98.94
DBNGM	99.08	98.76	99.34	99.57

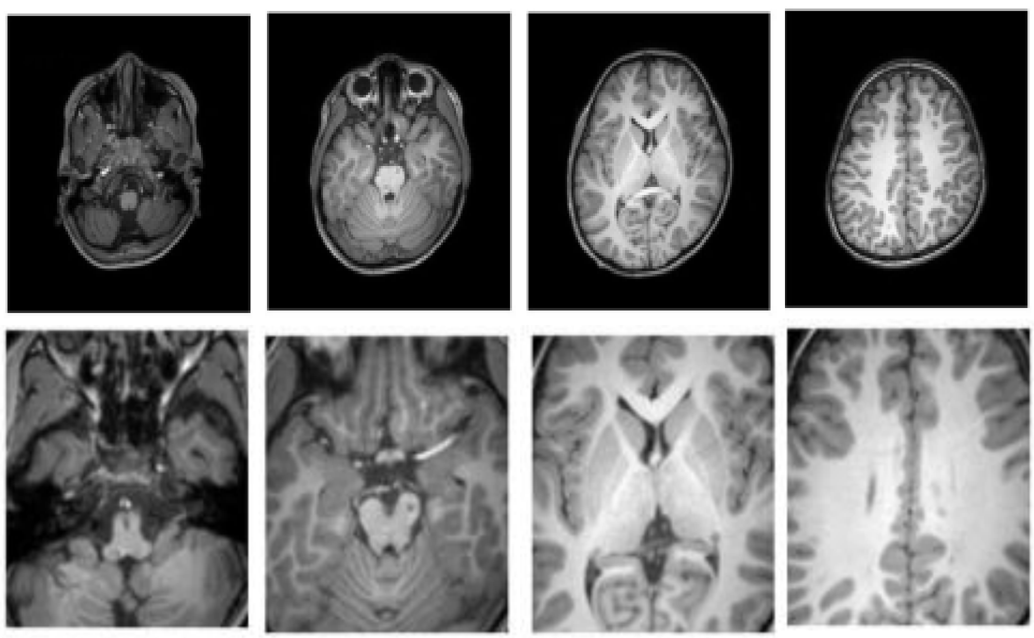


Figure 3 features extracted by dragon fly

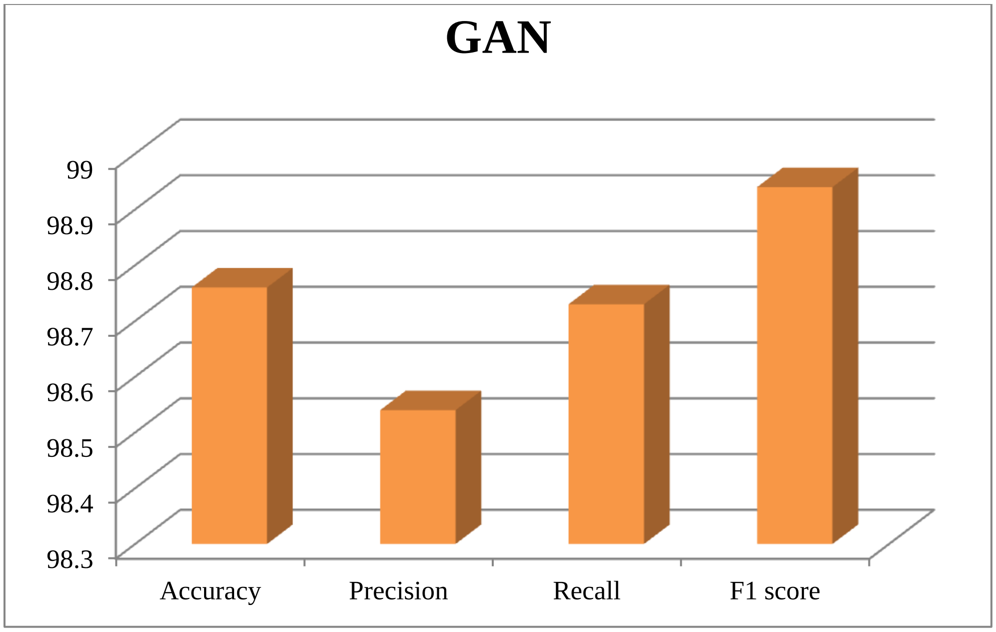


Figure 4. Performance of GAN

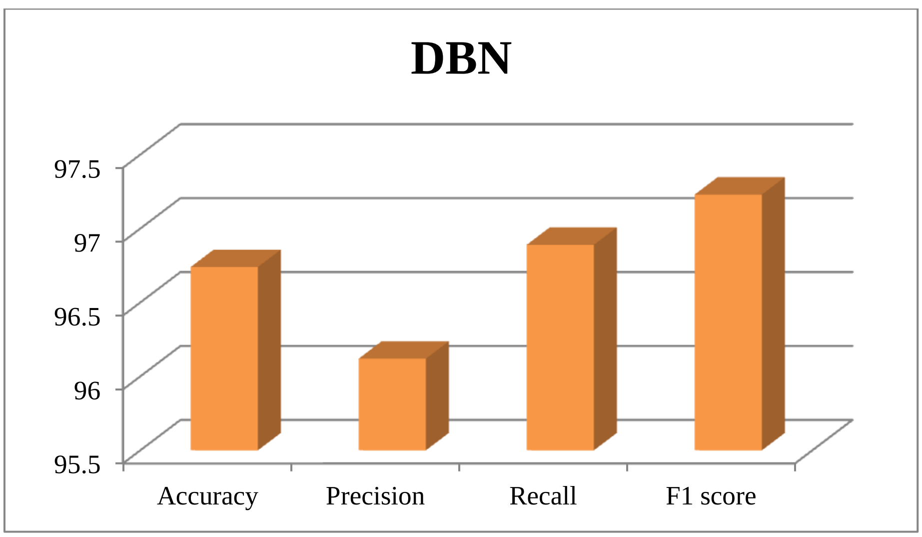


Figure 5 Evaluation metrics graph for DBN

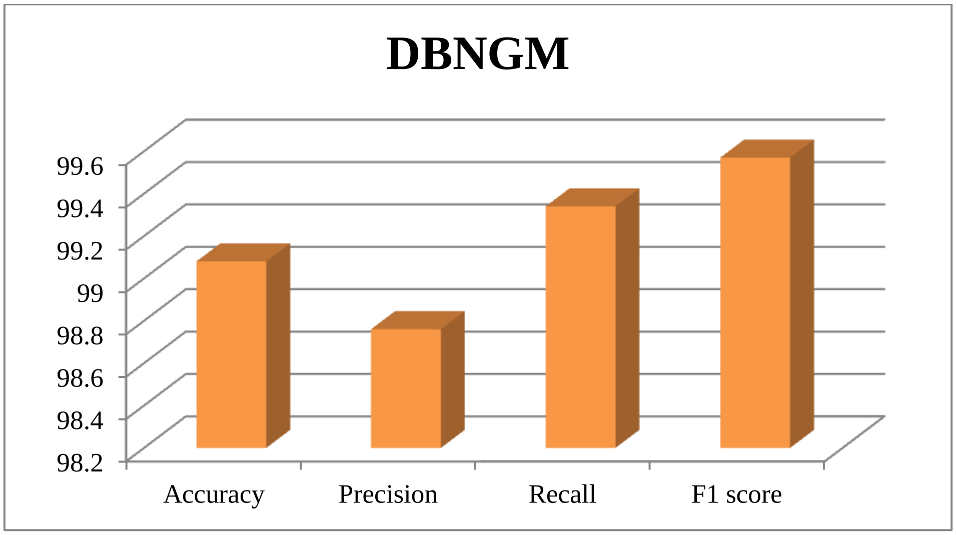


Figure 6 Evaluation metrics graph for Proposed DBNGM

The Precision averages of GAN, DBN and DBNGM are 96.12%, 98.54% and 98.76%. The Recall averages of GAN, DBN and DBNGM are 96.89%, 98.73% and 99.34%. The F-Score averages of GAN, DBN and DBNGM are 97.23%, 98.94%, and 99.57%. The accuracy averages of GAN, DBN and DBNGM are 96.74%, 98.76% and 99.08%.

V. DISCUSSION

By using DBNGM, the network can consistently learn significant information from the input sequence and process data more quickly. Without prior coordination, data were gathered for ABIDE I from a number of imaging centers. One major challenge for connectivity-based classifiers is the uncontrolled fluctuation in data caused by the disarray among imaging centers. Table 1 illustrates how the proposed deep neural network topologies dealt with the data heterogeneities in a satisfactory way and showed how merging several atlases and multisite data might yield useful discriminative information for autism diagnosis. A deep neural network may take into account several viewpoints to make up for the information lost from one when the input's heterogeneity rises.

VI. CONCLUSION

The study suggests utilizing fMRI data to identify autism spectrum disorders (ASD) using a deep learning-based intelligence system that uses the deep belief network generative model. The efficacy of the suggested strategy in the ASD identification task is demonstrated by experimental findings on the ABIDE dataset. This technique opens the door for the automatic diagnosis of ASD using discriminative picture markers. Additionally, our suggested method offers several advantages over some state-of-the-art techniques and achieves the best accuracy for ASD identification, suggesting that it has promise for the automatic diagnosis of ASD in clinical practice. The study intends to grow to incorporate a larger dataset and look into alternative deep learning architectures in future research. Improved medical decision-making and patient outcomes could result from the application of the suggested strategy in other medical specialties where diagnosis is difficult.

Reference

1. Bennett, M., Webster, A.A., Goodall, E. and Rowland, S., 2019. Life on the autism spectrum: Translating myths and misconceptions into positive futures. Springer.
2. Hyman, S.L., Levy, S.E., Myers, S.M., Kuo, D.Z., Apkon, S., Davidson, L.F., Ellerbeck, K.A., Foster, J.E., Noritz, G.H., Leppert, M.O.C. and Saunders, B.S., 2020. Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics*, *145*(1).
3. Chen, X., Cheng, G., Zou, D., Zhong, B. and Xie, H., 2023. Artificial Intelligent Robots for Precision Education. Educational Technology & Society, 26(1), pp.171-186.
4. Bauskis, A., Strange, C., Molster, C. and Fisher, C., 2022. The diagnostic odyssey: insights from parents of children living with an undiagnosed condition. *Orphanet Journal of Rare Diseases*, *17*(1), pp.1-13.
5. Salari, N., Rasoulpoor, S., Rasoulpoor, S., Shohaimi, S., Jafarpour, S., Abdoli, N., Khaledi-Paveh, B. and Mohammadi, M., 2022. The global prevalence of autism spectrum disorder: a comprehensive systematic review and meta-analysis. *Italian Journal of Pediatrics*, *48*(1), pp.1-16.
6. Shehab, M., Abualigah, L., Shambour, Q., Abu-Hashem, M.A., Shambour, M.K.Y., Alsalibi, A.I. and Gandomi, A.H., 2022. Machine learning in medical applications: A review of state-of-the-art methods. *Computers in Biology and Medicine*, *145*, p.105458.
7. Arabahmadi, M., Farahbakhsh, R. and Rezazadeh, J., 2022. Deep learning for smart Healthcare—A survey on brain tumor detection from medical imaging. *Sensors*, *22*(5), p.1960.
8. Yufei Wang, Jianxin Wang, Fang-Xiang Wu, Rahmatjan Hayrat, and Jin Liu. Aimafe: autism spectrum disorder identification with multi-atlas deep feature representation and ensemble learning. Journal of Neuroscience Methods, 343:108840, 2020.
9. Y. Ma et al., "Multi-Scale Dynamic Graph Learning for Brain Disorder Detection with Functional MRI," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, pp. 3501-3512, 2023.
10. K. Devika, D. Mahapatra, R. Subramanian and V. R. M. Oruganti, "Outlier-Based Autism Detection Using Longitudinal Structural MRI," in IEEE Access, vol. 10, pp. 27794-27808, 2022.

11. Z. Sherkatghanad et al., "Automated detection of autism spectrum disorder using a convolutional neural network," Frontiers in neuroscience, vol. 13, p. 1325, 2020.
12. J. Shin *et al*., "Identifying ADHD for Children With Coexisting ASD From fNIRs Signals Using Deep Learning Approach," in *IEEE Access*, vol. 11, pp. 82794-82801, 2023.
13. N. Rusli, S. N. Sidek, H. M. Yusof, N. I. Ishak, M. Khalid and A. A. A. Dzulkarnain, "Implementation of Wavelet Analysis on Thermal Images for Affective States Recognition of Children With Autism Spectrum Disorder," in IEEE Access, vol. 8, pp. 120818-120834, 2020.
14. S. Jain, H. K. Tripathy, S. Mallik, H. Qin, Y. Shaalan and K. Shaalan, "Autism Detection of MRI Brain Images Using Hybrid Deep CNN With DM-Resnet Classifier," in IEEE Access, vol. 11, pp. 117741-117751, 2023.
15. C. Craddock, Y. Benhajali, C. Chu, F. Chouinard, A. Evans, A. Jakab, B. S. Khundrakpam, J. D. Lewis, Q. Li, M. Milham, C. Yan, and P. Bellec, “The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives,” Neuroinformatics, 2013.
16. S. Mirjalili, "Dragonfly algorithm: a new metaheuristic optimization technique for solving singleobjective, discrete, and multi-objective problems," Neural Computing and Applications, vol. 27, pp. 1053- 1073, 2016.
17. C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," ACM SIGGRAPH computer graphics, vol. 21, pp. 25-34, 1987.

18. Mohamed, A.R., Dahl, G. and Hinton, G., 2009, December. Deep belief networks for phone recognition. In *Nips workshop on deep learning for speech recognition and related applications* (Vol. 1, No. 9, p. 39).
19. Li, T., Zhang, J. and Zhang, Y., 2014, October. Classification of hyperspectral image based on deep belief networks. In *2014 IEEE international conference on image processing (ICIP)* (pp. 5132-5136). IEEE