IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 1, February 2026, pp. 810~823
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i1.pp810-823

Neuro-DANet: dual attention deep neural network long short-term memory for autism spectrum disorder detection

Sujatha Hanumantharayappa, Manjula Rudragouda Bharamagoudra
School of Electronics and Communication Engineering, REVA University, Bengaluru, India

Article Info

Article history:

Received Jun 26, 2025
Revised Dec 25, 2025
Accepted Jan 10, 2026

Keywords:

Autism spectrum disorder
Continuous wavelet transforms
Dual-attention deep neural network
Functional magnetic resonance imaging
Long short-term memory
Principal component analysis

ABSTRACT

Autism spectrum disorder (ASD) is neurological illness affects ability of individuals to communicate and interact socially, and it is diagnosed in any time. Early detection of ASD is especially significant due to its subtle characteristics and high costs associated with the detection process. Traditional deep learning (DL) models struggle to capture intricate spatiotemporal dependencies in functional magnetic resonance imaging (fMRI) data, resulting in minimized detection performance and poor generalization. To address these drawbacks, the proposed Neuro-DANet combines a dual-attention deep neural network (DA-DNN) with long short-term memory (LSTM) to efficiently learn spatial and temporal features from fMRI scans. The continuous wavelet transform (CWT) is used to extract multi-scale features and the principal component analysis (PCA) is utilized to dimensionality reduction, which enhances robustness and efficacy. The dual self-attention mechanism improves the interpretability of the model by focusing on critical brain regions and time steps that are most relevant to ASD severity. The developed Neuro-DANet obtains the highest accuracy of 98.51% on autism brain imaging data exchange (ABIDE)-I and 98.81% on ABIDE-II datasets when compared with traditional algorithms.

Corresponding Author:

Sujatha Hanumantharayappa
School of Electronics and Communication Engineering, REVA University
Bengaluru, India
Email: sujag12@gmail.com

1. INTRODUCTION

Autism spectrum disorder (ASD) is neurodevelopmental illness considered with essential difficulties with social communication, a restricted range of interests, repetitive behaviors and atypical perceptual responses [1]–[3]. Symptoms of ASD generally emerge on early childhood, although social deficits are not noticeable in difficult social environments [4], [5]. The increasing prevalence of ASD has made it an essential public health concern [6]. Early and accurate diagnosis of ASD is significant for enabling timely interventions, which improve quality of life for individuals with the disorder [7]. Though the heterogeneous nature of ASD, with its wide range of severity and symptom levels, makes diagnosis challenging, cause misdiagnosis and delays [8]. Moreover, ASD diagnosis is challenging process which includes series of careful steps that involve long-term clinical monitoring, early assessment through caregiver and professional interviews with the physician [9], [10]. Recently, machine learning (ML) and deep learning (DL) models have been implemented to enhance ASD diagnosis [11]. These algorithms have been employed for neuroimaging information, especially functional magnetic resonance imaging (fMRI) and structural magnetic resonance imaging (sMRI) to detecting patterns which differentiate individuals with ASD from neurotypical controls [12], [13]. Recently, DL algorithms are preferred over conventional ML algorithms because of their

capability for automatically learning difficult hierarchical features from raw data, particularly in high-dimensional areas like neuroimaging [14]. Unlike ML algorithms, which rely on handcrafted features, deep models extract rich, multi-scale representations directly from fMRI and sMRI scans [15]. This is essential in autism diagnosis, where the subtle spatial and temporal patterns are complex to capture [16]. Self-attention mechanism improves the DL models through focusing on relevant brain regions, enhancing interpretability and accuracy [17]. This enables a network for capturing long-range dependencies without the drawbacks of sequences. Feature extraction by wavelet transforms (WT) enhances robustness and highlights the ASD specific abnormalities much efficiently [18]. The dimensionality reduction through principal component analysis (PCA) refines these features for effective model training [19]. These models improve a much precise, scalable and automatic algorithm for ASD from brain scans.

Song *et al.* [20] presented a novel diagnosis algorithm that combined graph convolutional networks (GCN) with dual transformer architectures, optimized by co-training strategy. Initially, a transformer was determined for capturing intricate temporal features from fMRI data, which were essential to understanding brain activity over time. The next transformer was applied to improve fusion of temporal features with spatial features learned through GCN, efficiently integrating dimensions of neuroimaging data. A co-training strategy was introduced to simultaneously use fMRI and sMRI data, enhancing the capacity of the model for generalization across various datasets. The fMRI images include redundant features that hinder learning efficacy and lose critical data. Tang *et al.* [21] suggested the graph neural network (GNN) and long short-term memory (LSTM) for ASD. Suggested model captured spatial attributes in fMRI information through GNN and aggregated temporal data of dynamic functional connectivity by LSTM for generating much complete spatio-temporal feature representation of fMRI. The dynamic graph pooling algorithm was developed for extracting the last node representation from the dynamic graph representation. To address variable dependencies on dynamic feature connectivity in time scales, method introduced jump connection mechanism for improving data extraction among internal units and captured attributes in various time scales. The model struggles with imbalanced data, leading to biased learning and poor generalization. Liu *et al.* [22] developed multi-atlas deep ensemble network (MADE) for ASD, which combined multi atlases of fMRI information by weighted deep ensemble network. The developed model combined demographic data into a prediction workflow that improves diagnosis of ASD performance and provides much holistic perspective in patient profiling. The model failed to extract spatial and temporal dynamics in fMRI data, which degrades detection performance. Ashraf *et al.* [23] introduced a 57-layer convolutional neural network (CNN) architecture named NeuroNet57, which extracted features from factually of fMRI. After pre-training on brain tumor data, introduced method was able to extract female phenotypic features from the autism brain image. Then, used an ant colony enabled system to select a feature subset, reducing extracted features size. However, introduced CNN model tends to lose fine-grained spatial features in its initial layers.

Khan and Katarya [24] implemented a bat algorithm-particle swarm optimization-LSTM (BAT-PSO-LSTM) network for ASD diagnosis. Here, utilized three different distinct datasets such as adults, children and toddlers, for comprehensive analysis of algorithms. BAT and PSO select the features and feed them to the adaptive feature fusion technique and LSTM classifier. The implemented model mitigated challenges like overfitting, slow training, model interpretability, generalization ability, and minimized training time. The implemented deep model causes overfitting in small neuroimaging data, which minimizes generalization ability of the model. Sriramakrishnan *et al.* [25] developed fractional whale-driving training-based optimization (FWCTBO) with CNN-enabled transfer learning (TL) to detect ASD. Developed model was designed through including fractional calculus (FC), whale optimization algorithm (WOA), and driving training-based optimization (DTBO), which trained CNN-TL hyperparameters. Additionally, CNN used hyperparameters from the trained methods such as ShuffleNet and AlexNet. For enhancing detection efficacy, the nub area was identified and processed using a functional connectivity-enabled whale driving training optimization (WDTBO) approach. Developed model failed to focus on much informative data, which degrades the detection performance. Traditional DL-based algorithms struggle to completely extract complex spatio-temporal dependencies that exist in fMRI data, limiting their detection accuracy for ASD. It struggles with feature redundancy; loss of spatial structure and overfitting issues integrated with high-dimensional neuroimaging data. Moreover, class imbalance in ASD severity level minimizes the model's robustness and generalization ability. The existing algorithms don't combine global spatial patterns and temporal dynamics. The main aim of this manuscript is to develop a DL framework that correctly detects ASD by fMRI data. The proposed method aims to efficiently capture spatiotemporal patterns by a combination of a self-attention mechanism and an LSTM architecture. This process improves feature learning by WT for multi-resolution analysis and PCA for dimensionality reduction. This model mitigates the issues like feature redundancy, spatial data loss and class imbalance, detection performance and generalization ability. The primary contributions of the manuscript are described as follows:

i) Developed a novel neurodevelopmental dual attention LSTM network (Neuro-DANet) model which combines dual-attention deep neural network (DNN) with LSTM for capturing spatial and temporal dependences in fMRI data, which enhances ASD severity detection and severity classification through learning deep and context-aware features.
ii) Employed continuous wavelet transform (CWT) for capturing multi-scale features from fMRI images, which captures global and fine-grained patterns. This improves the sensitivity of the model for subtle ASD-related abnormalities.
iii) Included the self-attention mechanism for focus on primary brain areas and time sequences, enhancing feature relevance and detection clarity. This attention mechanism highlights many informative patterns for ASD detection and severity classification.

This research manuscript is systemized as follows: section 2 provides a detailed explanation of the proposed model. Section 3 gives the outcomes and comparison of a proposed model. Section 4 concludes a manuscript.

2. RESEARCH METHOD

In this article developed the Neuro-DANet, which integrates self-attention DNN and LSTM for ASD detection. The datasets used for this manuscript are autism brain imaging data exchange (ABIDE)-I and ABIDE-II, then the images are pre-processed by using image resizing and data augmentation. The meaningful features are extracted by using WT and dimensionality is minimized by using PCA. At last, the self-attention DNN-LSTM network is used to detect the different classes of ASD. Figure 1 represents the process of ASD severity classification using self-attention DNN-LSTM.

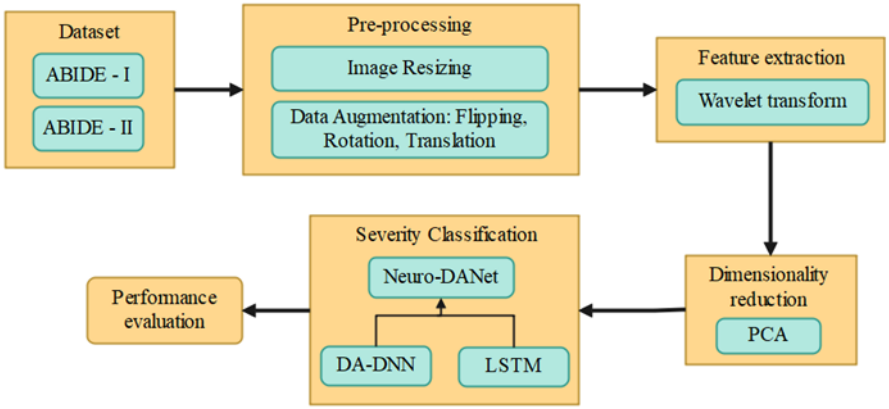

Figure 1. Process of ASD severity classification using self-attention DNN-LSTM

2.1. Dataset

In this article, used ABIDE dataset much claimed for its extensive neuroimaging data, which contains ABIDE-I [26] and ABIDE-II [27]. Dataset includes fMRI scans from various participants. The ABIDE-I dataset contains 419 individuals diagnosed to ASD and 530 neurotypical controls. ABIDE-II dataset contains 92 ASD and 103 neurotypical controls. This substantial participant tool, sourced from multiple international research facilities, improves statistical robustness and supports wide generalization ability. Table 1 represents the dataset description of ABIDE-I and Table 2 represents the dataset description of ABIDE-II dataset, and Figures 2 and 3 show the sample images of ABIDE-I and ABIDE-II datasets.

Table 1. Dataset description of ABIDE-I dataset

Classes	ASD	TD	Total images
Number of samples	1058	1163	2221

Table 2. Dataset description of ABIDE-II dataset

Classes	Miled	Moderate	Severe	TD	Total images
Number of samples	160	277	228	445	1110

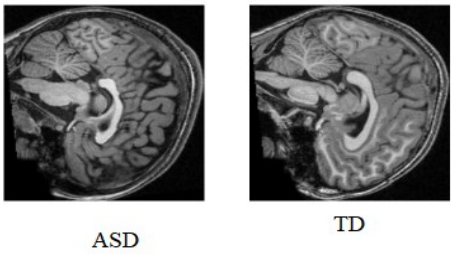

Figure 2. Sample images of ABIDE-I dataset

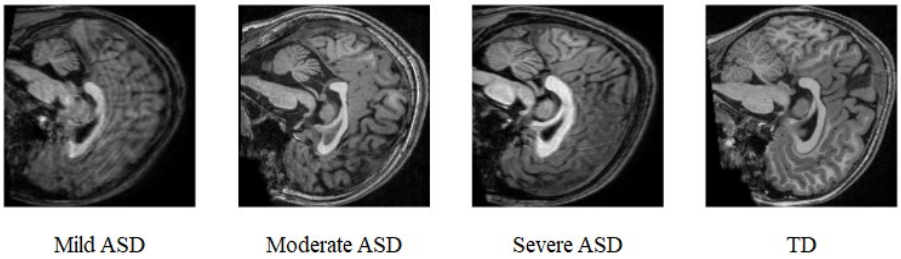

Figure 3. Sample images of ABIDE-II dataset

2.2. Pre-processing

Initially, for ABIDE-II dataset, split the data based on the social responsiveness scale (SRS) T-value by standard annotations. ASD Severity is classified into 4 groups for each domain depending on the SRS value. SRS total T-score provided in ABIDE-II phenotypic data is applied to fMRI subjects, due to SRS values are clinical or behavioral measures, not based on imaging modality. The SRS values for every individual class is described as follows:

- Typically developing (TD)–SRS value ≤59
- Mild ASD–SRS value range 60-65
- Moderate ASD–SRS value range 66-75
- Severe ASD–SRS value ≥76

2.2.1. Image resizing

In this article, the images from fMRI scans are resized to a fixed dimension of 224×224. This resizing is essential because neuroimaging data generally varies in resolution due to various scanners and positions of the brain slice. The DL models require uniform input dimensions for effective batch processing and training stability.

2.2.2. Data augmentation

In this article, data augmentation techniques like flipping, rotation, and translation are utilized for simulating variation in MRI acquisition without changing the diagnostic value of images. These techniques are employed for enhancing generalization ability of the model. These translations simulate general variations in MRI acquisition like hemispheric symmetry, head positioning that alters diagnostic data in fMRI slices. Augmentation helps the model to learn spatially robust features and overcomes the effects of class imbalance in ASD severity data.

- Flipping: it refers to mirroring images with axis like horizontal and vertical flipping. Flipping doesn't distort clinical features and helps the model to learn invariant features on both sides of the brain.
- Rotation: it helps the model to learn rotation-invariant features and minimizes overfitting for alignment patterns.
- Translation: it shifts the whole image horizontally, vertically by a few pixels or millimeters. This ensures a model for recognizing features regardless of their accurate spatial position, enhancing robustness for minor spatial shifts.

2.3. Feature extraction

WT provides the capability for decomposing fMRI images into multiple frequency and resolution levels, enabling that to extract both global structure patterns and fine-grained information. Traditional CNN models like MobileNet, ResNet learn the features by stacked convolutions without explicitly separating the frequency scales, which limits its sensitivity for subtle textures, especially in initial layers. Additionally, WT is inherently appropriate for small data because of its handcrafted and domain-aware development, where deep CNN models generally over-parameterized for limited fMRI data and leads to overfitting. Initially, time-frequency elements are extracted in every signal through the CWT. Coefficient of CWT is determined as signal convolution $x(t)$ with a translated and scaled version of wavelet $\psi_{a,b}(t)$ which is given as (1).

$$CWT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty\infty} x(t).\psi^* \left(\frac{t-b}{a}\right) dt \quad (1)$$

In (1), the a is wavelet scale, the b is time shift location and $*$ is difficult conjugate. A Morlet wavelet is chosen as a mother wavelet which has a better proportion (1.03) among frequency band and wavelet scale, that supports for interpreting outcomes in a frequency domain. Through varying a wavelet scale and translation with a localized time index, an image is generated that represents the amplitude of features across different scales. Amplitude is varied with time for generating scalogram images, that is, accurate value of WT coefficients. WT extracts spatial and frequency data in multiple resolutions, allowing for capturing meaningful patterns from fMRI data. This helps to highlight subtle brain activity differences like ASD through preserving essential structural and functional information while minimizing noise. Trade-off of WT:

– Lower levels (1-2) retain high-frequency noise and micro-patterns, that do not generalize well.
– Higher levels (5-6) over-smoothing images, losing essential anatomical information.

In this article, level 4 as the optimum balance, which preserves fine to midrange information and global brain structure. It efficiently balances the information retention and minimizes noise, essential to differentiate subtle ASD variations. From the WT, a total 55,906 features are extracted and given to the dimensionality reduction phase. Figures 4 and 5 represent the extracted features for ABIDE-I and ABIDE-II datasets.

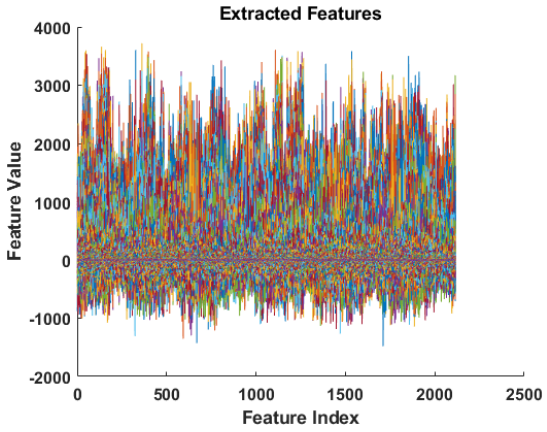

Figure 4. Extracted features for ABIDE-I dataset

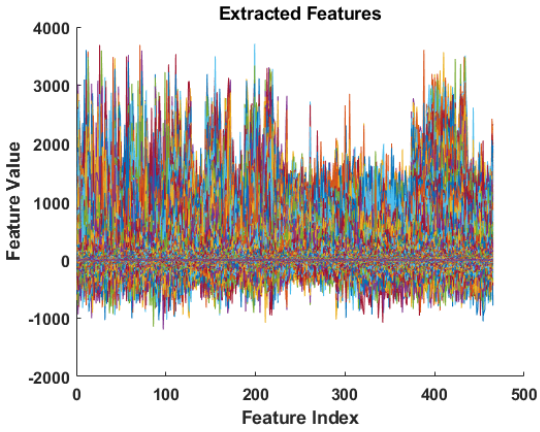

Figure 5. Extracted features for ABIDE-II dataset

2.4. Principal component analysis for dimensionality reduction

PCA preserves the internal architecture of wavelet-transformed fMRI features through capturing principal axes of variance that reflect much informative and correlated spatial-frequency patterns. This model ensures the primary statistical structure of brain abnormalities integrated with the severity of ASD that remains intact in a minimized dimensional feature space, facilitating effective detection. PCA identifies the principal components and the directions in feature space that extract many variances. However, Chi-square and ANOVA are suitable for categorical features, whereas here the features are continuous values. Figures 6 and 7 represent the transformed features using PCA on ABIDE-I and ABIDE-II datasets.

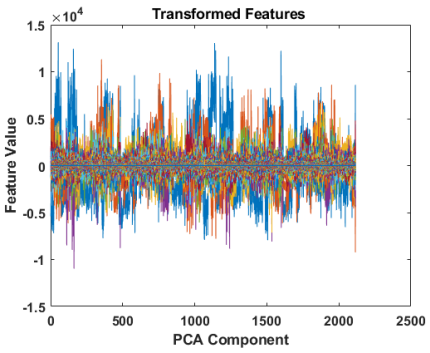

Figure 6. Transformed features for ABIDE-I dataset

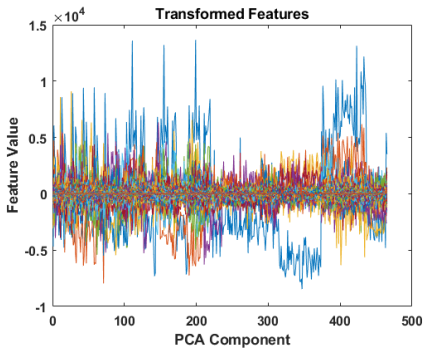

Figure 7. Transformed features for ABIDE-II dataset

2.5. Classification

Self-attention mechanism is efficient for analysis of fMRI because that extracts long-range spatial and temporal relationships, without relying on a sequential process. This process allows fast training and good scaling with huge fMRI data. In this article, self-attention in dual attention blocks focuses on essential brain areas and time points, improving feature representation while maintaining training stability by residual connections. An additional post-addition self-attention layer refines these features before feeding them into LSTM, which models the temporal dynamics. This integrated algorithm enhances the model's capability for identifying subtle and complex brain patterns, causing much precise fMRI detection when compared with traditional attention mechanisms. Figure 8 represents the architecture of Neuro-DANet model for ASD.

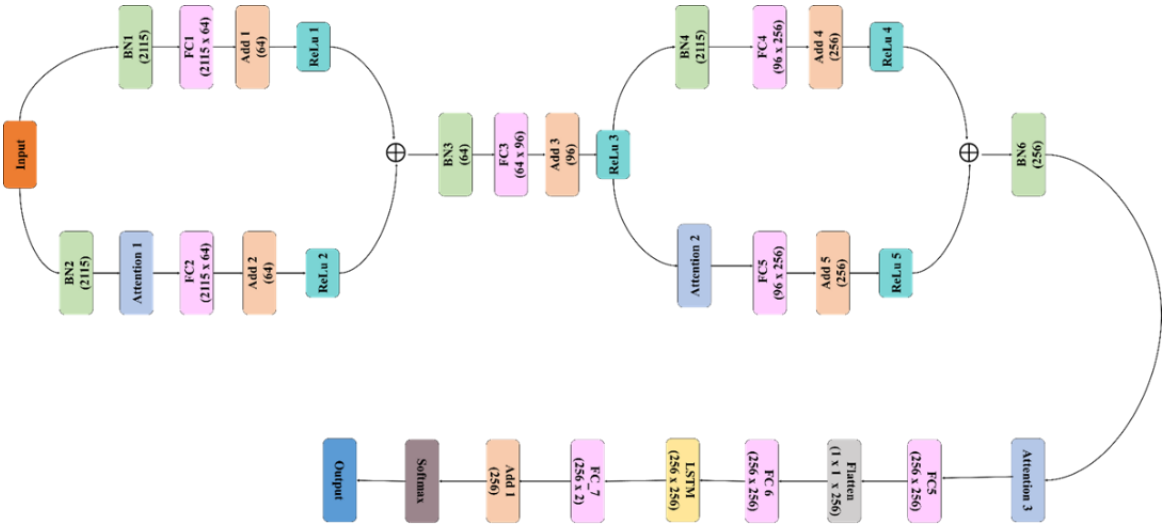

Figure 8. Architecture of Neuro-DANet model for ASD detection

2.5.1. Deep neural network

DNN is class of artificial neural network (ANN) includes multiple processing layers, able to learn difficult features and patterns from data. Layers in a DNN are classified into input, hidden, and output layers, with every layer being fully connected. This architecture enabled a method for processing and analyzing data at different abstraction levels, effectively improving their predictive abilities. DNN is trained by forward propagation and back propagation approaches, and gradient descent is utilized for optimizing the weights. In process of forward propagation, every neuron receives inputs from the prior layer, which is transmitted by a weighted sum, followed by non-linear activation function for generating the result. The mathematical expression for this process is given in (2).

$$z^{[l]} = \sigma\left(W^{[l]}\alpha^{[l-1]} + b^{[l]}\right) \quad (2)$$

For a given layer l, their result is represented as $z^{[l]}$ which is executed by employing σ an activation function is employed on the weighted sum input and bias term for that layer. Here, weighted input is acquired through multiplying the result from prior layer $\alpha^{[l-1]}$ with weight matrix $W^{[l]}$ in the present layer. The $b^{[l]}$ represents a bias term, which is introduced for additional adjustment, and facilitates model performance optimisation. After acquiring the gradient, gradient descent is utilized for updating the weights of the network. The mathematical expression for the weight update formula is given in (3). In the (3), the α represents learning rate.

$$W_{new} = W_{old} - \alpha \frac{\partial L}{\partial W} \tag{3}$$

2.5.2. Self-attention mechanism

For global dependencies of input sentence, two significant reasons that need to be mitigated: i) measuring attention of every word in an input sequence and ii) extracting sentence sequence data. For capturing global dependencies of images in data, measure the attention of every image in the data at first. Self-attention is much scalable and parallel attention measuring technique. Self-attention is considered as content-enabled query process, that executes attention function on group of queries and packages that to matrix Q. In similar time, keys and values are filled to matrices K and V. In an area of image processing, self-attention mechanism is utilized for dependency of every word with entire signals. For obtaining modality correlation, the dot product is computed. The matrix resulting from a multiplication of Q and K^T is a relationship among each image and whole other images. Correlation value of every modality is produced through softmax function. At last, this value is multiplied with mapping matrix V of S for acquiring words representation to extract global dependence data. Self-attention mechanism is executing the attention weight into every modality, and its mathematical expression is given as (4). In the (4), the $H^t = \{h_1, h_2, \dots, h_m\} \in R^{n \times dim_h}$ and $H = \{h_1, h_2, \dots, h_n\} \in R^{n \times dim_h}$ represent a hidden representation. Next, $Q, K,$ and V represents mapped matrices, that is initialized with multiplying input embedding and respective weight matrix and its mathematical expression is given as (5) to (7). In (5) to (7), the $Q, K, V \in R^{n \times 2dim_w}$ represents mappings of segment representation with data $H = \{\tilde{h}_1, \tilde{h}_2, \dots, \tilde{h}_n\} \in R^{n \times dim_h}$, the $w^Q, w^K \; and \; w^V$ represents learnable parameter matrices. Calculate its attention values for acquiring self-attention representation and its mathematical expression is given as (8). In (8), the $\sqrt{d_k}$ represents a scaling factor and that dimension is setted to a hidden representation.

$$\tilde{h}_i = (h_i + H^t), \;\; i \in [1, n] \tag{4}$$

$$Q = \tilde{H} w^Q \tag{5}$$

$$K = \tilde{H} w^K \tag{6}$$

$$V = \tilde{H} w^V \tag{7}$$

$$Z = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) V \tag{8}$$

2.5.3. Long short-term memory

LSTM is variant of recurrent neural network (RNN), for time series data, which is much suitable to capture different data features because of additional storage units able to store historical data. LSTM network architecture employed three gates such as forget, input and output gates in every LSTM neuron. The LSTM network process better when the input features are independent of each other. The uncorrelation nature of PCA-transformed features helps the LSTM model to focus on learning temporal relationships in data instead of handling the correlation among features. This process makes the training process much stable and effective. The f_t represents the rest of the information on prior LSTM neuron after fed to forget gate, the i_t represents amount of information for input on present LSTM neuron, the $\tilde{c}_t$ represents candidate state cell value and the c_t represents result amount of data on LSTM neuron. Mathematical expression for every gate of the LSTM cell is given as from (9) to (14).

$$f_t = \sigma\left(W_{xf} \cdot x_t + W_{hf} \cdot h_{t-1} + b_f\right) \tag{9}$$

$$i_t = \sigma(W_{xi} \cdot x_t + W_{hi} \cdot h_{t-1} + b_i) \tag{10}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (11)$$

$$\tilde{c}_t = tanh(W_{xc} \cdot x_t + W_{hc} \odot h_{t-1} + b_c) \quad (12)$$

$$o_t = \sigma(W_{xo} \cdot x_t + W_{ho} \cdot h_{t-1} + b_o) \quad (13)$$

$$h_t = o_t \odot tanh(c_t) \quad (14)$$

In (9) to (14), the σ represents sigmoid activation function, a $\odot$ represents Hadamard product and x_t represents present input vector. The $W_{xf}, W_{xi}, W_{xc}, W_{xo}$ represents respective weight vector for input, the h_{t-1} represents a state vector of a prior cell unit's hidden layer. The $W_{hf}, W_{hi}, W_{hc}, W_{ho}$ represents hidden layer weights on state vector, the b_f, b_i, b_c, b_o represents bias vector. The proposed method efficiently captures spatiotemporal patterns by a combination of a self-attention mechanism and an LSTM architecture. This process improves feature learning by WT for multi-resolution analysis and PCA for dimensionality reduction. This model mitigates the issues like feature redundancy, spatial data loss and class imbalance, detection performance and generalization ability.

3. RESULTS AND DISCUSSION

The developed model is simulated on MATLAB 2020 R and used system configuration are i5 processor, 8 GB RAM, and Windows 10 (64 bit). The validation metrics such as accuracy, sensitivity, specificity, precision, F1-score, and Matthew's correlation coefficient (MCC) are considered in this article to validate the performance. The parameters of the model are described in Table 3. Figures 9 and 10 represent the tested sample images using ABIDE-I and ABIDE-II datasets.

Table 3. Model parameters

Parameters	Value
Max epochs	100
Initial learning rate	0.01
L2 regularization	1.00E-04
Learning rate drop factor	0.2
Batch size	16
Learn rate schedule	Piecewise
Learn rate drop period	100
Gradient decay factor	0.91

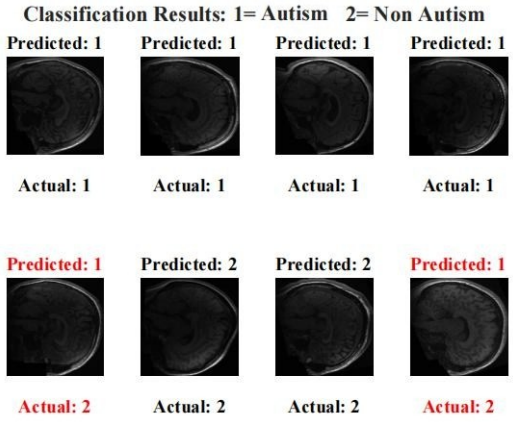

Figure 9. Tested sample images of ABIDE–I dataset

Table 4 represents an evaluation of the developed Neuro-DANet with different classifiers on ABIDE-I and ABIDE-II datasets. Different classifiers, such as gated recurrent unit (GRU), bidirectional long short-term memory (Bi-LSTM), DNN, and conventional LSTM, are taken to validate a performance of developed model. The proposed method aims to efficiently capture spatiotemporal patterns by a combination of a self-attention mechanism and an LSTM architecture. This process improves feature

learning by WT for multi-resolution analysis and PCA for dimensionality reduction. This model mitigates the issues like feature redundancy, spatial data loss and class imbalance, detection performance and generalization ability. Developed Neuro-DANet obtains the highest accuracy of 98.51% on ABIDE-I and 98.81% on ABIDE-II dataset.

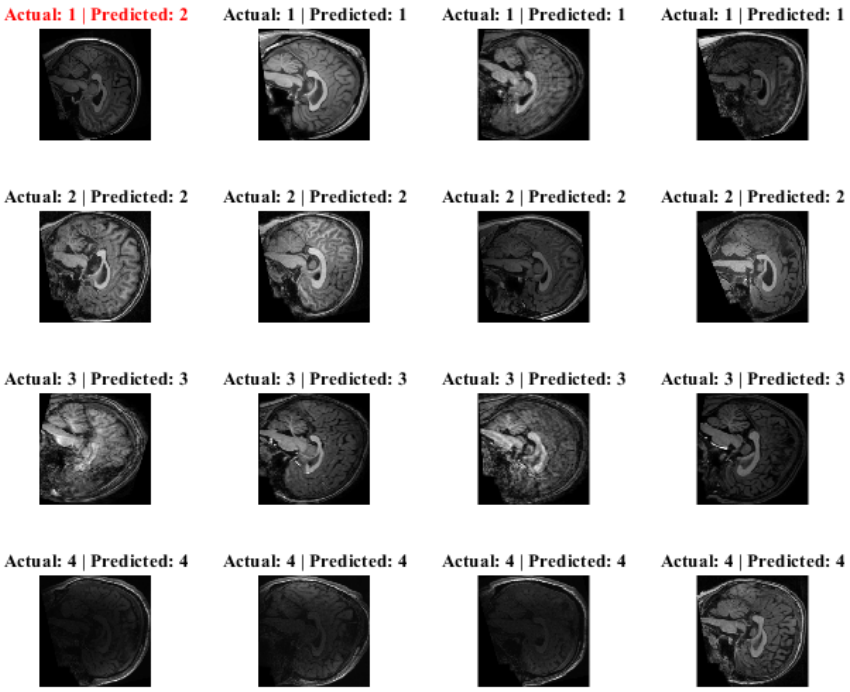

Figure 10. Tested sample images of ABIDE-II dataset

Table 4. Evaluation of developed Neuro-DANet with different classifiers on ABIDE-I and ABIDE-II datasets

Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	MCC (%)
			ABIDE-I			
GRU	76.64	74.35	81.07	79.70	76.93	71.59
Bi-LSTM	81.65	79.03	83.19	82.63	80.79	78.19
DNN	79.26	79.21	75.69	76.83	78.00	74.66
LSTM	95.28	94.44	93.48	92.77	93.60	91.84
Neuro-DANet	98.51	98.11	98.11	98.73	98.42	96.24
			ABIDE-II			
GRU	96.43	95.99	96.86	95.12	95.55	93.64
Bi-LSTM	94.24	93.84	93.20	93.09	93.47	91.00
DNN	97.19	96.44	97.06	97.13	96.78	95.71
LSTM	98.05	97.31	97.73	97.40	97.36	96.60
Neuro-DANet	98.81	98.81	99.60	98.86	98.83	98.43

Table 5 represents the evaluation of 5-fold cross-validation for the developed Neuro-DANet with different classifiers on ABIDE-I and ABIDE-II datasets. The different classifiers such as GRU, Bi-LSTM, DNN, and conventional LSTM, are taken to evaluate a performance of a developed model. The proposed method aims to efficiently capture spatiotemporal patterns by a combination of a self-attention mechanism and an LSTM architecture. This process improves feature learning by WT for multi-resolution analysis and PCA for dimensionality reduction. This model mitigates the issues like feature redundancy, spatial data loss and class imbalance, detection performance and generalization ability. On 5-fold cross validation, the developed model obtains a high accuracy of 98.51% on ABIDE-I and 98.81% on ABIDE-II datasets.

Table 6 represents the evaluation of different k-fold validation sets on ABIDE-I and ABIDE-II datasets. K-fold validation sets such as K=3, K=5, K=7 and K=8 is evaluated for the developed Neuro-DANet. In that, the developed model shows the highest accuracy on K=5 when compared with other k-fold validation sets.

Table 5. Evaluation of 5-fold cross validation for developed Neuro-DANet with different classifiers using ABIDE-I and ABIDE-II datasets

Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	MCC (%)
			ABIDE-I			
GRU	76.07	73.38	80.10	79.14	76.15	71.02
Bi-LSTM	80.68	78.06	82.62	81.66	79.82	77.62
DNN	78.29	78.64	75.12	75.01	76.78	74.09
LSTM	94.32	93.87	92.51	91.80	92.82	91.28
Neuro-DANet	98.51	98.11	98.11	98.73	98.42	96.24
			ABIDE-II			
GRU	95.43	94.99	95.86	94.20	94.59	93.04
Bi-LSTM	93.24	92.84	93.98	93.00	92.92	90.00
DNN	96.19	95.44	96.96	96.93	96.18	95.01
LSTM	97.05	98.31	97.03	97.00	97.65	96.00
Neuro-DANet	98.81	98.81	99.60	98.86	98.83	98.43

Table 6. Evaluation of different k-fold validation sets for the developed Neuro-DANet using ABIDE-I and ABIDE-II datasets

K-fold validation	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	MCC (%)
			ABIDE-I			
K=3	96.90	95.64	95.08	96.00	95.82	94.08
K=5	98.51	98.11	98.11	98.73	98.42	96.24
K=7	97.50	96.46	96.54	96.62	96.54	95.40
K=8	95.20	92.55	92.13	93.95	93.25	92.70
			ABIDE-II			
K=3	96.61	96.36	95.45	95.42	95.89	94.58
K=5	98.81	98.81	99.60	98.86	98.83	98.43
K=7	95.56	94.63	95.09	94.65	94.64	93.43
K=8	94.12	93.64	93.42	94.88	94.26	93.45

Table 7 represents the evaluation of different L2 regularization values for the developed Neuro-DANet on ABIDE-I and ABIDE-II datasets. On L2 regularization values of 1.00E-04, the developed Neuro-DANet obtained the highest accuracy of 98.51% on ABIDE-I and 98.81% on ABIDE-II datasets. Table 8 represents the evaluation of different batch sizes for the developed Neuro-DANet using ABIDE-I and ABIDE-II datasets. On Batch size of 16, developed Neuro-DANet obtained the highest accuracy of 98.51% on ABIDE-I and 98.81% on ABIDE-II datasets.

Table 7. Evaluation of different L2 regularization values of the developed Neuro-DANet on ABIDE-I and ABIDE-II datasets

L2 regularization	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	MCC (%)
			ABIDE-I			
4.00E-04	92.94	92.63	91.33	92.76	92.70	92.28
3.00E-04	96.52	93.67	95.94	94.03	93.85	94.86
2.00E-04	97.62	97.25	98.84	97.02	97.13	95.05
1.00E-04	98.51	98.11	98.11	98.73	98.42	96.24
			ABIDE-II			
4.00E-04	94.38	93.42	93.36	92.50	92.96	91.26
3.00E-04	95.58	94.00	94.68	93.27	93.63	92.78
2.00E-04	97.24	95.17	96.21	95.29	95.23	95.69
1.00E-04	98.81	98.81	99.60	98.86	98.83	98.43

Table 8. Evaluation of different batch sizes of developed Neuro-DANet using ABIDE-I andABIDE-II datasets

Batch size	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	MCC (%)
			ABIDE-I			
128	92.68	91.56	90.09	91.71	91.63	90.92
64	95.61	95.68	94.36	94.41	95.04	93.13
32	96.97	96.65	94.07	95.76	96.20	94.32
16	98.51	98.11	98.11	98.73	98.42	96.24
			ABIDE-II			
128	95.25	94.74	94.22	93.65	94.20	92.21
64	94.01	93.74	94.87	93.50	93.62	91.85
32	97.10	96.84	95.56	95.77	96.31	94.55
16	98.81	98.81	99.60	98.86	98.83	98.43

Table 9 represents the evaluation of different initial learning rates for the developed Neuro-DANet using ABIDE-I and ABIDE-II datasets. A learning rate of 0.01, the developed Neuro-DANet obtained the highest accuracy of 98.51% on ABIDE-I and 98.81% on ABIDE-II datasets. Table 10 represents an evaluation of statistical analysis for the developed Neuro-DANet using ABIDE-I and ABIDE-II datasets. Evaluation metrics such as t-statistic, p-value and degrees of freedom are considered to evaluate statistical performance of a developed model. Figures 11 and 12 represent confusion matrix of ABIDE-I and ABIDE-II datasets respectively. Figures 13 and 14 represent ROC curve of ABIDE-I and ABIDE-II datasets respectively.

Table 9. Evaluation of different initial learning rate for the developed Neuro-DANet using ABIDE-I and ABIDE-II datasets

Initial learning rate	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	MCC (%)
			ABIDE-I			
0.04	95.81	94.40	94.59	95.78	95.09	93.14
0.03	94.27	95.24	94.75	93.80	94.51	92.25
0.02	97.17	96.90	95.45	96.51	96.71	95.60
0.01	98.51	98.11	98.11	98.73	98.42	96.24
			ABIDE-II			
0.04	95.77	94.55	94.81	95.83	95.19	93.59
0.03	94.66	93.20	93.91	94.24	93.72	93.53
0.02	96.94	96.99	94.10	93.47	95.20	97.29
0.01	98.81	98.81	99.60	98.86	98.83	98.43

Table 10. Statistical analysis of the developed Neuro-DANet model using ABIDE-I and ABIDE-II datasets

Datasets	t-statistic	p-value	Degrees of freedom
ABIDE-I	4.6057	0.0037	6
ABIDE-II	2.5538	0.0433	6

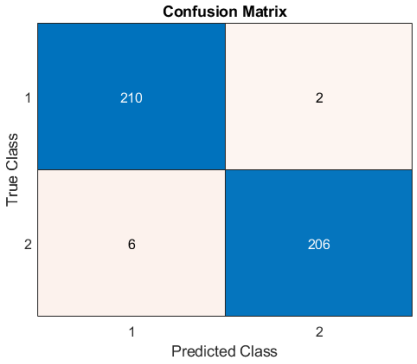

Figure 11. Confusion matrix on ABIDE-I dataset

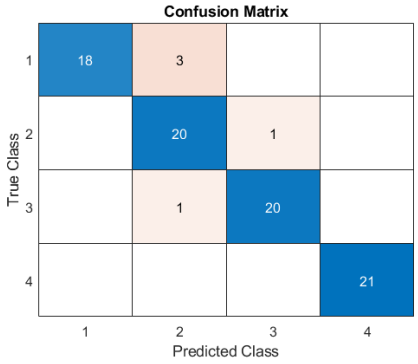

Figure 12. Confusion matrix on ABIDE-II dataset

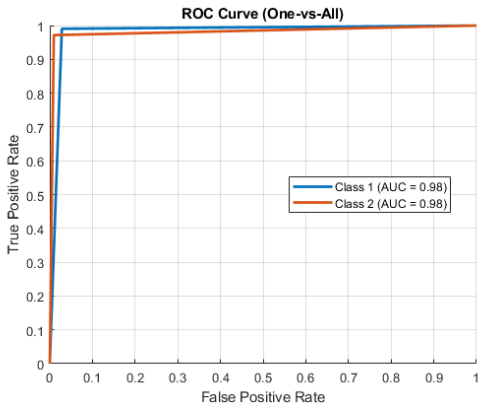

Figure 13. ROC curve on ABIDE-I dataset

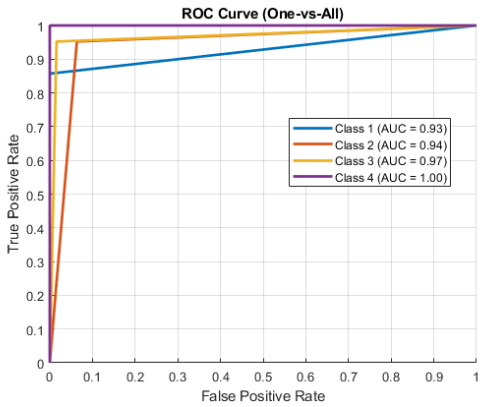

Figure 14. ROC curve on ABIDE-II dataset

3.1. Comparative analysis

In this section, the developed Neuro-DANet is compared to existing algorithms like GCN+dual transformer [20], GNN+LSTM [21], MADE for ASD [22] and CNN-based NeuroNet [23] using ABIDE-I dataset. The existing algorithms like GCN+dual transformer [20], GNN+LSTM [21], and CNN-based NeuroNet [23] are compared using ABIDE-II dataset. The proposed method aims to efficiently capture spatiotemporal patterns by a combination of a self-attention mechanism and an LSTM architecture. This process improves feature learning by WT for multi-resolution analysis and PCA for dimensionality reduction. This model mitigates the issues like feature redundancy, spatial data loss and class imbalance, detection performance and generalization ability. From Table 11, the developed Neuro-DANet model obtained the highest accuracy of 98.51% on ABIDE-I and 98.81% on ABIDE-II datasets when compared to existing algorithms.

Table 11. Comparative analysis of developed Neuro-DANet model with existing algorithms on ABIDE-I and ABIDE-II datasets

Datasets	Methods	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)
ABIDE-I	GCN+dual transformer [20]	79.47	78.97	82.11	NA
	GNN+LSTM [21]	80.4	NA	NA	NA
	MADE for ASD [22]	75.20	NA	82.90	69.70
	CNN-based NeuroNet [23]	92.21	NA	NA	NA
	Proposed Neuro-DANet	98.51	98.73	98.11	98.11
ABIDE-II	GCN+dual transformer [20]	76.55	76.20	77.14	NA
	GNN+LSTM [21]	79.63	NA	NA	NA
	CNN-based NeuroNet [23]	93.49	NA	NA	NA
	Proposed Neuro-DANet	98.81	98.86	98.81	99.60

3.2. Research implications

This research produced a number of main contributions that are relevant to the development of neuroimaging-based ASD detection methods, including:

i) Advanced ASD detection modelling: demonstrates how combining dual-attention mechanisms with LSTM enhances model capability to detect ASD severity from difficult fMRI data.
ii) Improves understanding of brain dynamics: offers insights into spatiotemporal brain activity patterns relevant to ASD, helps in neuroscientific studies of functional brain abnormalities.
iii) Multi-resolution feature learning: validates the efficacy of wavelet-based multi-scale feature extraction in capturing global structures and subtle variations in neuroimaging data.
iv) Enhances model interpretability: highlights the attention mechanism in making neural networks more transparent and clinically interpretable for detection.
v) Helps broad dataset generalization: addresses the challenges of dataset variability and class imbalance, paving the way for more robust AI applications in heterogeneous clinical settings.

4. CONCLUSION

This manuscript developed Neuro-DANet framework that combined dual-attention DNN and LSTM for efficient ASD severity detection using fMRI data. By using wavelet-based approach to capture meaningful multi-scale features and then the dimensionality is reduced by using the PCA approach. The developed model addressed challenges like feature redundancy, loss of spatial-temporal dependencies and class imbalance. By incorporating a self-attention mechanism, the method improves interpretability and enables focused attention on the most relevant brain regions and time sequences. The proposed method aims to efficiently capture spatiotemporal patterns by a combination of a self-attention mechanism and an LSTM architecture. This process improves feature learning by WT for multi-resolution analysis and PCA for dimensionality reduction. This model mitigates the issues like feature redundancy, spatial data loss and class imbalance, detection performance and generalization ability. The experimental evaluation is assessed on ABIDE-I and ABIDE-II datasets determine that Neuro-DANet obtains superior accuracy and generalization ability compared to existing algorithms, showing its robustness and scalability for automatic ASD detection. The developed Neuro-DANet demonstrates high robustness and accuracy in ASD severity detection, though there are certain areas to focus on for future work. Combining multimodal data like sMRI to further improve the detection accuracy and generalization.

ACKNOWLEDGMENTS

The authors would like to thank REVA University and REVA AICTE IDEA Lab for the continuous support, infrastructure and encouragement to carry out the research.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Sujatha Hanumantharayappa	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓
Manjula Rudragouda Bharamagoudra		✓				✓	✓	✓	✓	✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available in ABIDE-I and ABIDE-II at:

– https://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html, reference number [26].
– https://fcon_1000.projects.nitrc.org/indi/abide/abide_II.html, reference number [27].

REFERENCES

[1] A. Weerasekera, A. I.-Mărgineanu, G. P. Nolan, and M. Mody, "Subcortical-cortical white matter connectivity in adults with autism spectrum disorder and schizophrenia patients," *Psychiatry Research: Neuroimaging*, vol. 340, 2024, doi: 10.1016/j.pscychresns.2024.111806.

[2] C. Rakshe, S. Kunneth, S. Sundaram, M. Murugappan, and J. F. A. Ronickom, "Autism spectrum disorder diagnosis using fractal and non-fractal-based functional connectivity analysis and machine learning methods," *Neural Computing and Applications*, vol. 36, no. 20, pp. 12565–12585, 2024, doi: 10.1007/s00521-024-09770-3.

[3] Y. Wang *et al.*, "Abnormalities in cerebellar subregions' volume and cerebellocerebral structural covariance in autism spectrum disorder," *Autism Research*, vol. 18, no. 1, pp. 83–97, 2025, doi: 10.1002/aur.3287.

[4] U. Singh, S. Shukla, and M. M. Gore, "Detection of autism spectrum disorder using multi-scale enhanced graph convolutional network," *Cognitive Computation Systems*, vol. 6, no. 1–3, pp. 12–25, 2024, doi: 10.1049/ccs2.12108.

[5] V. Jain, C. T. Rakshe, S. S. Sengar, M. Murugappan, and J. F. A. Ronickom, "Age-and severity-specific deep learning models for autism spectrum disorder classification using functional connectivity measures," *Arabian Journal for Science and Engineering*, vol. 49, no. 5, pp. 6847–6865, 2024, doi: 10.1007/s13369-023-08560-8.

[6] D. Dhar, M. Chaturvedi, S. Sehwag, C. Malhotra, C. Saraf, and M. Chakrabarty, "Gray matter volume correlates of co-occurring depression in autism spectrum disorder," *Journal of Autism and Developmental Disorders*, vol. 54, no. 10, 2024, doi: 10.1007/s10803-024-06602-0.

[7] D. D. Jemima, A. G. Selvarani, and J. D. L. Lovenia, "Multi-view united transformer block of graph attention network based autism spectrum disorder recognition," *Frontiers in Psychiatry*, vol. 16, 2025, doi: 10.3389/fpsyt.2025.1485286.

[8] C. Li *et al.*, "Gray matter asymmetry alterations in children and adolescents with comorbid autism spectrum disorder and attention-deficit/hyperactivity disorder," *European Child and Adolescent Psychiatry*, vol. 33, no. 8, pp. 2593–2604, 2023, doi: 10.1007/s00787-023-02323-4.

[9] J. S. Zhu, Q. Gong, M. T. Zhao, and Y. Jiao, "Atypical brain network topology of the triple network and cortico-subcortical network in autism spectrum disorder," *Neuroscience*, vol. 564, pp. 21–30, 2025, doi: 10.1016/j.neuroscience.2024.11.034.

[10] N. E. H. Mezrioui, K. Aloui, A. N.-Ali, and M. S. Naceur, "Automatic characterization of cerebral MRI images for the detection of autism spectrum disorders," *Intelligence-Based Medicine*, vol. 9, 2024, doi: 10.1016/j.ibmed.2023.100127.

[11] R. Ghnemat, N. Al-Madi, and M. Awad, "An intelligent approach for autism spectrum disorder diagnosis and rehabilitation features identification," *Neural Computing and Applications*, vol. 37, no. 4, pp. 2557–2580, 2025, doi: 10.1007/s00521-024-10770-6.

[12] M. I. Nazir, T. Mazumder, M. M. Islam, M. A. Ehsan, R. Rahman, and T. Helaly, "Enhancing autism spectrum disorder diagnosis through a novel 1D CNN-based deep learning classifier," in *2024 3rd International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)*, 2024, pp. 1–6, doi: 10.1109/ICAEEE62219.2024.10561645.

[13] J. Gao and S. Song, "A hierarchical feature extraction and multimodal deep feature integration-based model for autism spectrum disorder identification," *IEEE Journal of Biomedical and Health Informatics*, vol. 29, no. 7, pp. 4920-4931, 2025, doi: 10.1109/JBHI.2025.3540894.

[14] L. Kang, M. Chen, J. Huang, and J. Xu, "Identifying autism spectrum disorder based on machine learning for multi-site fMRI," *Journal of Neuroscience Methods*, vol. 416, 2025, doi: 10.1016/j.jneumeth.2025.110379.
[15] Z. Guan *et al.*, "Dynamic graph transformer network via dual-view connectivity for autism spectrum disorder identification," *Computers in Biology and Medicine*, vol. 174, 2024, doi: 10.1016/j.compbiomed.2024.108415.
[16] Y. Wang, H. Long, T. Bo, and J. Zheng, "Residual graph transformer for autism spectrum disorder prediction," *Computer Methods and Programs in Biomedicine*, vol. 247, 2024, doi: 10.1016/j.cmpb.2024.108065.
[17] L. Ruan *et al.*, "Brain functional gradient and structure features in adolescent and adult autism spectrum disorders," *Human Brain Mapping*, vol. 45, no. 11, 2024, doi: 10.1002/hbm.26792.
[18] P. Qing *et al.*, "Structure-function coupling in white matter uncovers the hypoconnectivity in autism spectrum disorder," *Molecular Autism*, vol. 15, no. 1, 2024, doi: 10.1186/s13229-024-00620-6.
[19] X. Shan *et al.*, "Disentangling the individual-shared and individual-specific subspace of altered brain functional connectivity in autism spectrum disorder," *Biological Psychiatry*, vol. 95, no. 9, pp. 870–880, 2024, doi: 10.1016/j.biopsych.2023.09.012.
[20] T. Song, Z. Ren, J. Zhang, Y. Qu, Y. Cui, and Z. Liang, "Multimodal autism spectrum disorder method using GCN with dual transformers," *IEEE Access*, vol. 13, pp. 16324–16337, 2025, doi: 10.1109/ACCESS.2025.3532302.
[21] J. Tang, J. Chen, M. Hu, Y. Hu, Z. Zhang, and L. Xiao, "Diagnosis of autism spectrum disorder (ASD) by dynamic functional connectivity using GNN-LSTM," *Sensors*, vol. 25, no. 1, 2024, doi: 10.3390/s25010156.
[22] X. Liu, M. R. Hasan, T. Gedeon, and M. Z. Hossain, "MADE-for-ASD: a multi-atlas deep ensemble network for diagnosing autism spectrum disorder," *Computers in Biology and Medicine*, vol. 182, 2024, doi: 10.1016/j.compbiomed.2024.109083.
[23] A. Ashraf, Q. Zhao, W. H. Bangyal, M. Raza, and M. Iqbal, "Female autism categorization using CNN based NeuroNet57 and ant colony optimization," *Computers in Biology and Medicine*, vol. 189, 2025, doi: 10.1016/j.compbiomed.2025.109926.
[24] K. Khan and R. Katarya, "AFF-BPL: an adaptive feature fusion technique for the diagnosis of autism spectrum disorder using Bat-PSO-LSTM based framework," *Journal of Computational Science*, vol. 83, 2024, doi: 10.1016/j.jocs.2024.102447.
[25] G. V Sriramakrishnan, P. M. Paul, H. Gudimindla, and V. Rachapudi, "Fractional whale driving training-based optimization enabled transfer learning for detecting autism spectrum disorder," *Computational Biology and Chemistry*, vol. 113, 2024, doi: 10.1016/j.compbiolchem.2024.108200.
[26] A. D. Martino and S. Mostofsky, "Autism brain imaging data exchange I," *Child Main Institute.* [Online]. Available: https://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html.
[27] A. D. Martino and S. Mostofsky, "Autism brain imaging data exchange II," *Child Main Institute.* [Online]. Available: https://fcon_1000.projects.nitrc.org/indi/abide/abide_II.html

BIOGRAPHIES OF AUTHORS

Sujatha Hanumantharayappa is working as a lecturer in the Department of Electronics and Communication Engineering at Government Polytechnic Mulabagal Kolar and currently, she is deputed for a part-time Ph.D. at REVA University Bangalore. She completed her M.Tech. degree in Electronics and Communication at UVCE Bangalore affiliated with Bangalore University. She has more than 13+ years of teaching experience and her research areas are medical image processing and machine learning. Her publications include a peer-reviewed research article in a Scopus-indexed journal and conference papers published by Springer Nature and IEEE, a leading academic publisher. She can be contacted at email: sujag12@gmail.com.

Manjula Rudragouda Bharamagoudra has 20 years of teaching, research and administrative experience. Some of the journals where her research articles published are Elsevier, Springer, and Wiley, having good impact factors. She has published papers in peer reviewed national and international journals, reputed national and international conferences and book chapter/book. She has citations more than 465 and h-index of 11. Her areas of interest are communication, AI, internet of things, wireless sensor networks, underwater communication, and robotics. She can be contacted at email: manjula.rb@reva.edu.in.