JURNAL INFOTEL
VOL. 18, NO. 3, AUGUST 2026, PP. 531–554. DOI:10.20895/INFOTEL.V18I3.11761

RESEARCH ARTICLE

EEG-Based Autism Spectrum Disorder Classification Using TFSNet with Multiscala PCA and ICA Denoising

Yunidar Yunidar[1,*], Muhammad Mirza Rahmat[2], Fithria Fithria[3], Melinda Melinda[4], Siti Rusdiana[5], and Rosmawinda Rosmawinda[6]

[1,2,4,6]Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Indonesia
[3]Department of Family Health Nursing, Faculty of Nursing, Universitas Syiah Kuala, Indonesia
[5]Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Indonesia

*Corresponding email: yunidar@usk.ac.id

Received: July 13, 2026; Revised: October 13, 2026; Accepted: October 20, 2026.

Abstract: Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that requires objective and non-invasive methods for early identification. This study proposes an EEG-based ASD classification framework utilizing a *Time-Frequency Synergy Network* (TFSNet) at the feature level, integrated with multiscale denoising methods. EEG data were collected from 16 children (8 with ASD and 8 with typical development as a control group) using a 16-channel OpenBCI system. The signals were filtered using a 0.5–40 Hz Butterworth bandpass filter and subsequently denoised via *Multiscale Principal Component Analysis* (MSPCA) and *Multiscale Independent Component Analysis* (MSICA). The signals were segmented into 2-second epochs with a 50% overlap. *Power Spectral Density* (PSD) and *bandpower* features were processed using a dual-branch TFSNet, which incorporates convolutional feature extraction, attention-based refinement, *Transformer* encoding, and *Adaptive Cross-Domain Attention* (ACDA). Performance was evaluated using subject-based 8-fold cross-validation. The TFSNet with MSPCA achieved an average epoch-level accuracy of 98.60%, precision of 98.78%, recall of 98.41%, F1-score of 98.59%, specificity of 98.79%, and an AUC of 99.85%. Although MSICA yielded a higher signal-to-noise ratio (24.48 dB compared to 22.24 dB), its classification accuracy was lower (98.07%). Statistical tests indicate a significant performance difference, with the MSPCA method performing better (*paired t-test*, $p = 0.0073$; Wilcoxon test, $p = 0.0078$).

Keywords: ASD, EEG, MSPCA, MSICA, TFSNet

© Jurnal Infotel ISSN: 2085-3688; e-ISSN: 2460-0997

1 Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by impairments in social communication, restricted behavioral flexibility, and repetitive patterns of activity. The diagnosis of ASD is commonly conducted through behavioral observation and clinical assessment. Although these approaches are widely used in clinical practice, they can be time-consuming and subject to subjective interpretation. Therefore, recent studies have explored electroencephalography (EEG) as a non-invasive modality to support objective and data-driven ASD identification [1–4]. EEG provides high temporal resolution and can capture neural activity patterns that may differ between ASD and normal subjects, particularly in spectral and frequency-domain characteristics.

EEG-based ASD classification has been widely investigated using handcrafted features, machine learning, and deep learning approaches [5–8]. Among the commonly used EEG features, Power Spectral Density (PSD) and bandpower features remain important because they represent neural oscillatory activity within specific frequency bands. In addition, several studies have transformed EEG signals into time-frequency or image-like representations to improve classification performance, making EEG analysis increasingly relevant to pattern recognition and computational intelligence frameworks [9–11]. However, the reliability of EEG-based classification strongly depends on the quality of the recorded signal and the effectiveness of the preprocessing stage.

EEG signals are highly sensitive to artifacts originating from eye movements, muscle activity, cardiac interference, electrode instability, and environmental noise [12–14]. This issue becomes more critical in practical acquisition systems such as OpenBCI, where portable and low-cost recording conditions may introduce additional noise and signal variability. Therefore, denoising is an essential step before feature extraction and classification. Several methods, including Multiscale Principal Component Analysis (MSPCA) and Multiscale Independent Component Analysis (MSICA), have been applied to suppress artifacts and improve EEG signal quality [14, 15]. Nevertheless, selecting an effective denoising method for EEG-based ASD classification remains challenging because excessive denoising may remove useful neural information, while insufficient denoising may preserve irrelevant artifacts.

To address this issue, this study proposes a comparative framework to evaluate two EEG denoising methods: Butterworth bandpass filtering followed by either MSPCA or MSICA [15]. We evaluate both methods using the same classification framework distinguishing between ASD and normal subjects—based on EEG signals acquired via OpenBCI. The primary objective is to determine which denoising strategy yields the best performance in feature-level classification using the Time-Frequency Synergy Network (TFSNet). Unlike previous TFSNet studies [7]or research applying MSPCA and MSICA with other classification models [14,15], this study focuses on evaluating these two denoising methods within a dual-branch TFSNet framework that integrates PSD and bandpower features via an Adaptive Cross-Domain Attention (ACDA) mechanism. Thus, the study's contribution lies in comparing the effects of MSPCA and MSICA on EEG signal quality and ASD classification performance within a unified evaluation framework.

The proposed workflow begins with Butterworth bandpass filtering in the 0.5–40 Hz range to preserve EEG frequency components relevant to brain activity. Following filtering, one of three denoising methods is applied to reduce artifacts; the data are then analyzed via MSPCA and MSICA components using a feature-level TFSNet architecture comprising

PSD and bandpower branches. Representations from both branches are fused using Adaptive Cross-Domain Attention (ACDA), and the final fused features are classified using a Multilayer Perceptron (MLP) to generate class probabilities for ASD or normal status.

To obtain a realistic, leakage-aware evaluation, this study uses an 8-fold subject-wise cross-validation protocol. In this protocol, data splitting is performed at the subject level rather than the segment level. Therefore, all EEG segments from the same subject are assigned exclusively to either the training or the testing fold. This strategy prevents subject-dependent information leakage and provides a more reliable estimate of model generalization. Feature standardization is also performed using a StandardScaler fitted only on the training data within each fold.

The model performance is evaluated using accuracy, balanced accuracy, precision, recall, F1-score, specificity, area under the curve (AUC), and confusion matrix analysis. Furthermore, statistical testing is conducted to determine whether the performance differences among Butterworth with MSPCA and Butterworth with MSICA are significant. The main contributions of this study are summarized as follows. First, it provides a controlled comparison of two multiscale denoising methods, MSPCA and MSICA, for OpenBCI EEG-based ASD classification following Butterworth bandpass filtering. Second, it employs a dual-branch TFSNet framework that fuses PSD and bandpower features. Third, it adopts a subject-wise cross-validation protocol to reduce evaluation bias caused by subject-dependent factors.

2 Research Method

Figure 1 illustrates the overall workflow of the proposed study. EEG signals were collected from 16 subjects, including 8 children diagnosed with ASD and 8 neurotypical controls. Recordings were performed using 16-channel OpenBCI systems at a sampling rate of 250 Hz. Raw EEG signals were initially inspected visually to identify gross artifacts and ensure data integrity.

Preprocessing of EEG signals began with a zero-phase Butterworth bandpass filter with a frequency range of 0.5-40 Hz to retain physiologically relevant brain activity while attenuating low-frequency drifts and high-frequency noise. Following this, one of two multiscale denoising methods, Multiscale Independent Component Analysis (MS-ICA) or Multiscale Principal Component Analysis (MS-PCA), was applied to reduce residual artifacts, including eye movements, muscle activity, and environmental interference. The cleaned EEG dataset was then segmented into 2-second windows as a step towards feature extraction.

From each segment, Power Spectral Density (PSD) and bandpower features were calculated for all 16 channels. These features served as inputs to the proposed TFSNet architecture, which incorporates dual-branch processing for PSD and bandpower representations. The outputs of the two branches were fused using an Adaptive Cross-Domain Attention (ACDA) module, and classification probabilities for ASD versus Normal were generated using a fully connected Multilayer Perceptron (MLP). For model evaluation, an 8-fold subject-wise cross-validation protocol was implemented, in which 14 subjects were used for training and 2 subjects were used for testing in each fold. All epochs belonging to the same subject were assigned exclusively to either the training or testing partition within each fold, thereby preventing information leakage between subjects. Feature standardization was performed using a StandardScaler fitted only on the training data of each fold.

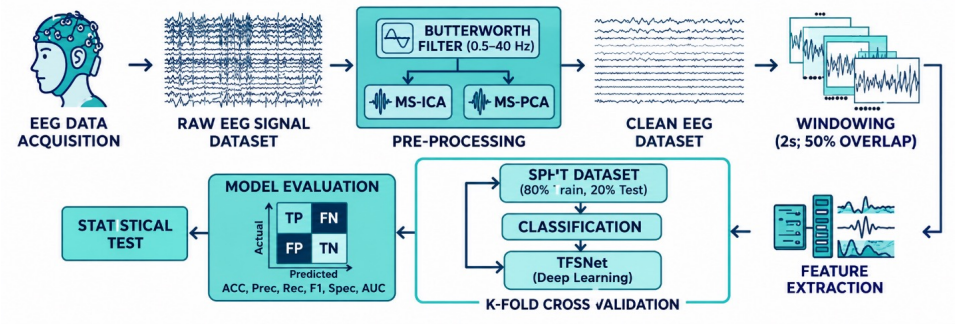

Figure 1: Workflow of EEG-based ASD classification using Butterworth bandpass filtering, MS-PCA, and MS-ICA with TFSNet.

The primary unit of evaluation in this study is the EEG epoch rather than the individual participant. During each subject-wise cross-validation fold, all epochs from the same subject were kept within the same partition, thereby preventing epochs from the same participant from appearing in both the training and testing sets. Consequently, the reported accuracy, precision, recall, F1-score, specificity, and AUC represent epoch-level classification performance. These metrics should not be interpreted as subject-level diagnostic accuracy, as multiple epochs can originate from the same participant. Subject-level aggregation was not performed in the current study and is therefore considered an important direction for future validation.

Classification performance was assessed using accuracy, balanced accuracy, precision, recall, F1-score, specificity, area under the receiver operating characteristic curve (AUC), and confusion matrix analysis. Statistical testing was conducted to determine whether TFSNet's classification performance differed significantly between the MSPCA and MSICA preprocessing approaches.

2.1 Dataset

The EEG dataset used in this study was acquired directly by the authors using an OpenBCI Cyton Board configured with 16 channels. Electrodes were placed according to the International 10-20 system, including Fp1, Fp2, F3, F4, F7, F8, C3, C4, T3, T4, T5, T6, P3, P4, O1, and O2. The recording setup also incorporated dedicated reference (REF) and ground (GND) electrodes integrated into the cap and positioned along the upper midline to ensure stable signal referencing and grounding. EEG signals were sampled at 250 Hz.

The participant group consisted of 16 children, divided into 8 individuals diagnosed with ASD and 8 neurotypical control participants. For ethical and privacy reasons, only group-level demographic information is reported. Data collection was conducted in a controlled environment in Banda Aceh, Indonesia, following international standard recording protocols. The study received ethical approval from the institutional Ethics Committee (Approval No. 117/EA/FK/2024) and adhered to the WHO guidelines (2011) [15, 16]. Each subject underwent two EEG recording sessions, each lasting 15 minutes. However, only one 15-minute resting-state recording session was selected for the preprocessing and

classification analysis. Using a 2-second sliding window with a step size of 1 second, this resulted in approximately 895 epochs per subject [17].

Due to the limited sample size, an 8-fold subject-wise cross-validation strategy was applied. In each fold, 14 subjects were used for training, and the remaining 2 for testing. Data splitting was performed at the subject level prior to segmentation to preserve subject labels and prevent leakage between training and testing datasets. All preprocessing parameters were estimated exclusively from the training subjects in each fold and subsequently applied to the corresponding test subjects. Additionally, model weights were reinitialized and trained from scratch for each fold to ensure unbiased, consistent performance evaluation.

2.2 Butterworth Band Pass Filter

The Butterworth band-pass filter is a standard preprocessing technique used to suppress frequency components outside a designated passband while preserving a maximally flat response within the passband. This band-pass characteristic is achieved by sequentially applying high-pass and low-pass filters, defined by the lower and upper cutoff frequencies, respectively.

In this study, EEG signals were processed using a fourth-order Butterworth band-pass filter with cutoff frequencies of 0.5–40 Hz. The lower cutoff attenuates slow-varying components, such as baseline drift and low-frequency physiological artifacts, while the upper cutoff attenuates high-frequency components outside the selected passband. The filter was implemented as an Infinite Impulse Response (IIR) filter and applied bidirectionally using forward–backward filtering to minimize phase distortion. This frequency range preserves EEG activity within the selected passband for subsequent multiscale denoising and feature extraction [14, 18]. The filter was implemented as an Infinite Impulse Response (IIR) structure and applied in a bidirectional (forward–backward) manner to minimize phase distortion and preserve the temporal integrity of the EEG signals. This filtering procedure was consistently applied to all channels before subsequent multiscale denoising and classification, following standard EEG preprocessing protocols reported in prior literature.

2.3 Multiscale Signal Decomposition Using Discrete Wavelet Transform

Discrete Wavelet Transform (DWT) is applied in this study as the initial multiscale decomposition stage for EEG signal preprocessing. This method is suitable for EEG analysis because EEG signals are generally non-stationary and contain time-varying spectral characteristics [19]. Through DWT, the EEG signal can be represented at different resolution levels, allowing low-frequency brain activity and high-frequency artifact components to be separated more effectively [14, 15, 20]. Therefore, DWT provides a time–frequency representation that is useful for subsequent denoising and component separation processes.

In the DWT framework, each EEG signal is passed through a pair of quadrature mirror filters, namely a low-pass filter and a high-pass filter. The low-pass filter produces approximation coefficients, which represent the low-frequency components of the signal, while the high-pass filter produces detail coefficients, which contain the high-frequency components. The approximation and detail coefficients are mathematically expressed as follows:

$$A[n] = \sum_{k} S[k] \cdot g[2n - k] \tag{1}$$

$$D[n] = \sum_{k} S[k] \cdot h[2n - k] \tag{2}$$

where $S[k]$ denotes the input EEG signal, $g[\cdot]$ represents the low-pass filter, and $h[\cdot]$ represents the high-pass filter. Based on Eq. (1) and Eq. (2), $A[n]$ contains the approximation information associated with low-frequency activity, whereas $D[n]$ contains the detail information associated with higher-frequency components. This separation enables the analysis of EEG artifacts that are dominant in specific frequency ranges.

The decomposition process is performed iteratively to obtain a multilevel signal representation. At each decomposition level, the approximation coefficients from the previous level are further decomposed into new approximation and detail coefficients. This process can be formulated as follows:

$$A_j[n] = \sum_{k} A_{j-1}[k] \cdot g[2n - k] \tag{3}$$

$$D_j[n] = \sum_{k} A_{j-1}[k] \cdot h[2n - k] \tag{4}$$

Based on Eq. (3) and Eq. (4), the approximation and detail coefficients represent the detail coefficients at the j-th decomposition level, respectively. The multilevel decomposition enables analysis of the EEG signal across multiple frequency scales, allowing noise and artifacts to be reduced without removing important neural information. This stage is particularly important because artifacts such as eye movements, muscle activity, and other non-neural interference may overlap with EEG rhythms in both time and frequency domains.

In this study, the Daubechies-4 (db4) wavelet was employed because it provides a good balance between time localization and frequency resolution for EEG signal analysis [14, 15]. The EEG signal was decomposed into three levels to obtain multiscale representations that preserve relevant brain activity while isolating unwanted components. The resulting wavelet coefficients at each scale were then processed using either Principal Component Analysis (PCA) or Independent Component Analysis (ICA), depending on the denoising framework being evaluated.

The overall framework of the proposed multiscale EEG denoising process is illustrated in Figure 2. The EEG signal is first decomposed using DWT into approximation and detail coefficients. The decomposed coefficients are then processed using either PCA in the MSPCA framework or ICA in the MSICA framework to suppress noise and artifact components. Finally, the corrected coefficients are reconstructed using inverse DWT to obtain the denoised EEG signal.

2.3.1 Multiscale Principal Component Analysis

In the Multiscale Principal Component Analysis (MSPCA) framework, the wavelet coefficients obtained from DWT are processed using PCA at each decomposition level. PCA is applied to transform the multichannel EEG coefficients into a set of uncorrelated components based on the variance distribution. Components with high variance are assumed to contain dominant EEG information, while components with low variance are considered less informative and may represent noise or artifact contributions.

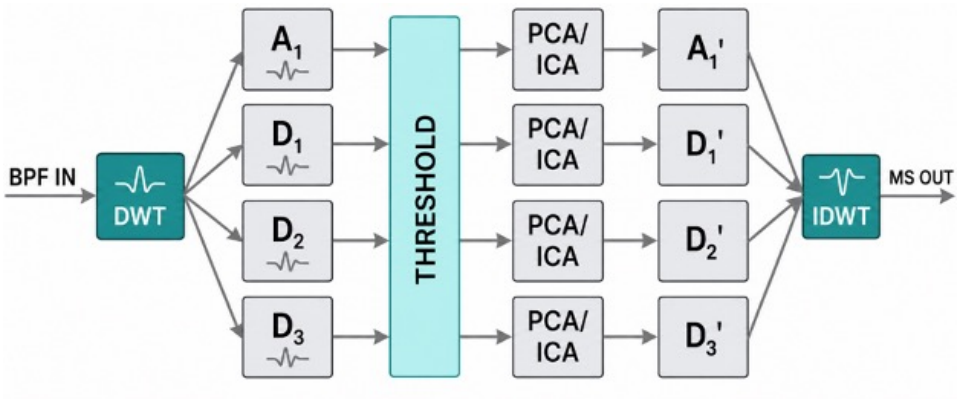

Figure 2: Structural diagram of multiscale EEG denoising using MSPCA and MSICA.

After component selection, the retained principal components are reconstructed as wavelet coefficients and then transformed back to the time domain using the inverse DWT. This process produces a denoised EEG signal while preserving the major signal structure. Therefore, MSPCA is used to reduce noise by combining multiscale decomposition and variance-based component selection without modifying the original mathematical structure of PCA [15,21].

2.3.2 Multiscale Independent Component Analysis

In the Multiscale Independent Component Analysis (MSICA) framework, ICA is applied to the wavelet coefficients obtained at each DWT decomposition scale. Unlike PCA, which separates components based on variance, ICA separates statistically independent sources contained in the EEG signal. This approach is suitable for EEG artifact removal because recorded EEG signals are mixtures of neural activity and non-neural artifacts, such as eye movements, muscle contractions, and environmental noise.

At each decomposition level, ICA separates the wavelet coefficients into independent components. Components associated with artifacts are attenuated or removed, while components representing neural activity are retained. The corrected wavelet coefficients are then reconstructed using inverse DWT to obtain the final denoised EEG signal. This multiscale strategy allows artifact suppression at different frequency resolutions, making MSICA more adaptive to non-stationary EEG signals [14,20].

In this study, both MSPCA and MSICA were implemented using the same DWT configuration to ensure a fair comparison. The db4 wavelet was selected as the mother wavelet, and three decomposition levels were used to capture relevant EEG frequency characteristics. For dimensionality reduction and source separation, 14 components were retained at each decomposition level. The ICA stage was implemented using the FastICA algorithm, while PCA was applied using standard principal component transformation. A fixed random seed of 42 was used throughout the decomposition and component analysis stages to ensure reproducibility. This configuration was selected to balance denoising performance, computational efficiency, and consistency across all EEG signal processing experiments.

2.4 Feature Extraction

Feature extraction is an important stage in the EEG-based classification pipeline because it transforms preprocessed EEG signals into numerical representations that can describe the spectral characteristics of brain activity. In this study, feature extraction was performed after the EEG signals were processed with a Butterworth band-pass filter (0.5-40 Hz) and denoised using one of the proposed denoising methods.

Two spectral feature groups were extracted from each EEG segment, namely Power Spectral Density (PSD) and bandpower features. PSD was used to describe the distribution of EEG signal power across frequency bins with a resolution of 1 Hz, while bandpower was used to summarize the total power within clinically relevant EEG frequency bands [18,22]. These two feature groups were then processed using a feature-level TFSNet architecture comprising a PSD branch and a bandpower branch. The representations learned from both branches were fused using Adaptive Cross-Domain Attention (ACDA), and the fused representation was classified using a Multilayer Perceptron (MLP) to produce ASD or normal class probabilities.

2.4.1 Power Spectral Density

Power Spectral Density (PSD) was used to estimate the distribution of EEG signal power over the frequency domain. In this study, PSD estimation was performed for each EEG channel and each segmented epoch after preprocessing and denoising. The PSD feature was selected because it is capable of representing the spectral characteristics of EEG signals and providing informative frequency-domain patterns to distinguish EEG activity in ASD from that in normal conditions. For a sampling frequency of 250 Hz, each segment consisted of 500 samples. PSD estimation was carried out using the Welch method, which divides each EEG segment into overlapping windows and averages the periodogram values to obtain a more stable spectral estimate [18,23]. A Hamming window was applied to each subsegment to reduce spectral leakage during frequency-domain transformation.

For an EEG segment with a length of samples, the periodogram of the -th window can be expressed as follows:

$$P_1(f) = \frac{1}{MU} \left| \sum_{n=0}^{M-1} x_l[n] w[n] e^{-j2\pi fn} \right|^2 \tag{5}$$

where $x_l[n]$ denotes the EEG signal in the l-th window, $w[n]$ represents the Hamming window, M is the number of samples in each window, and U is the normalization factor of the window function. The normalization factor is defined as:

$$U = \frac{1}{M} \sum_{n=0}^{M=1} w^2[n] \tag{6}$$

The final PSD estimate is obtained by averaging all periodograms, as formulated below:

$$S_x(f) = \frac{1}{L} \sum_{l=1}^{L} P_l \tag{7}$$

where L is the total number of windowed segments used in the Welch estimation. Based on Eq. (5), Eq. (6), and Eq. (7), the Welch method provides a stable PSD estimate by reducing variance through periodogram averaging.

In this study, the frequency range of 0.5–40 Hz was discretized into non-overlapping frequency bins with a resolution of 1 Hz. Therefore, 40 PSD features were obtained for each EEG channel. Since the EEG data comprised 16 channels, the PSD extraction yielded 640 PSD features per EEG segment. These PSD features were then used as the input of the PSD branch in the proposed TFSNet architecture.

2.4.2 Bandpower

Bandpower was extracted to represent the total EEG signal power within specific frequency bands. Unlike PSD, which provides detailed power information for each frequency bin, bandpower summarizes power across physiologically meaningful EEG rhythms. Therefore, bandpower features can provide a compact representation of brain activity and complement PSD-based representations.

In this study, bandpower was calculated by integrating the PSD values within predefined EEG frequency bands. The five frequency bands used in this study were delta, theta, alpha, beta, and gamma. These bands were defined as follows: delta, theta, alpha, beta, and gamma [22, 24].

Mathematically, the bandpower value for a frequency band $[f_1|f_2]$ can be expressed as:

$$P_{\text{band}} = \sum_{f=f_1}^{f_2} S_x(f)\Delta f \tag{8}$$

where $S_x(f)$ denotes the PSD estimate obtained using the Welch method, and Δf represents the frequency resolution. In this study, Δf was set to 1 Hz. Based on Eq. (8), the bandpower value was obtained by summing the PSD values within each predefined EEG frequency band.

In this study, five EEG frequency bands were considered: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–40 Hz). Bandpower was calculated by integrating the corresponding Power Spectral Density (PSD) values within each frequency band. For each of the 16 EEG channels, five bandpower features were extracted, resulting in 80 bandpower features per EEG segment. These features were used as inputs to the bandpower branch of the proposed TFSNet model. Unlike PSD features, which preserve detailed frequency-bin information, bandpower features summarize spectral activity within predefined frequency bands. Combining both representations enables TFSNet to learn complementary fine-grained spectral patterns and broader EEG rhythm characteristics for ASD and typically developing control classification.

2.5 Time-Frequency Synergy Network (TFSNet) Feature-Level Classifier

The proposed feature-level TFSNet classifier is designed to classify ASD and normal EEG signals using two complementary spectral feature groups, namely PSD-bin features and bandpower features. The overall structure of the proposed classifier is illustrated in Figure 3. The input EEG feature tensor consists of spectral features extracted from each EEG segment after preprocessing, denoising, and segmentation. These features are then divided into the PSD branch and the bandpower branch.

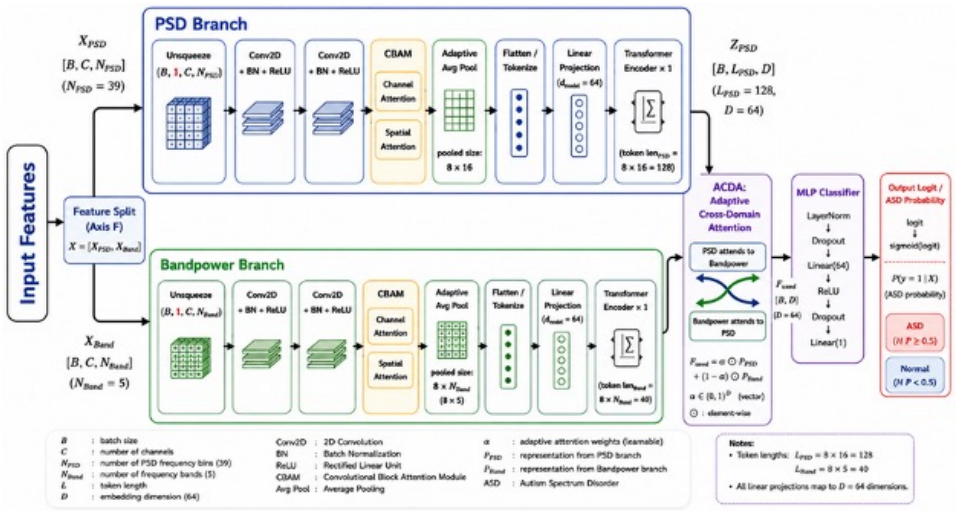

Figure 3: Flow diagram of the proposed feature-level TFSNet classifier.

The input feature tensor is represented as:

$$X \in \mathbb{R}^{B \times C \times F_{\text{total}}} \tag{9}$$

where B denotes the batch size, C represents the number of EEG channels, and F_total is the total number of extracted features. The input tensor is divided into PSD-bin and bandpower components as follows:

$$X = [X_{\text{PSD}}, X_{\text{Band}}] \tag{10}$$

where X_PSD contains 1-Hz PSD-bin features, while X_Band contains bandpower features from predefined EEG frequency bands. This separation allows the model to learn detailed frequency-bin information and compact rhythm-based EEG representations independently.

In each branch, the input feature is first reshaped into a two-dimensional feature map and then processed by convolutional layers, followed by batch normalization and ReLU activation. The convolutional layers are used to extract local relationships between EEG channels and spectral features. After the convolutional stage, the Convolutional Block Attention Module (CBAM) is applied to emphasize informative channel and spatial features while suppressing less relevant feature responses [7, 25, 26].

The refined feature maps are then reduced using adaptive average pooling and transformed into token sequences. These tokens are projected into a fixed embedding dimension and processed using a Transformer encoder to capture global dependencies among channel-feature representations. In general, the output representation of each branch can be written as:

$$Z_d = \text{Trans}(\text{Proj}(\text{Pool}(\text{CBAM}(\text{Conv}(d))))) + \text{PE}) \tag{11}$$

where $d \in \mathrm{PSD}, \mathrm{Band}$, PE denotes positional encoding, and Z_d represents the output feature representation from each branch. Through this process, the PSD branch produces Z_{PSD}, while the bandpower branch produces Z_{Band}.

After both branches generate their respective representations, Adaptive Cross-Domain Attention (ACDA) is applied to fuse information between PSD and bandpower domains. This mechanism enables the PSD representation to attend to bandpower information and, conversely, allows the bandpower representation to attend to PSD information. The attention operation is expressed as:

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{12}$$

where Q, K, and V denote query, key, and value matrices, respectively, and d_k is the dimension of the key vector. In this study, ACDA is used to model the interaction between detailed PSD-bin information and summarized bandpower characteristics [7, 25].

The two attended feature representations are then combined using an adaptive fusion coefficient. The final fused representation is defined as:

$$F_{\mathrm{fused}} = \alpha P_{\mathrm{PSD}} + (1 - \alpha)P_{\mathrm{Band}} \tag{13}$$

where P_{PSD} and P_{Band} are the pooled representations from the PSD and bandpower branches, respectively, and α is a learnable adaptive coefficient that controls the contribution of each feature domain. This fusion strategy enables the model to automatically determine the relative importance of PSD and bandpower features for each EEG segment.

Finally, the fused feature vector is passed into a Multilayer Perceptron (MLP) classifier. The output logit is converted into a class probability using sigmoid activation:

$$P(y - 1 \mid X) = \mathrm{sigmoid}(\mathrm{MLP}(F_{\mathrm{fused}})) \tag{14}$$

where $P(y = 1X)$ represents the probability that the EEG segment belongs to the ASD class. If the probability value exceeds the classification threshold, the segment is classified as ASD, otherwise, it is classified as normal. Overall, the proposed feature-level TFSNet integrates convolutional feature extraction, CBAM-based feature refinement, Transformer-based global modeling, and ACDA-based adaptive fusion to improve ASD EEG classification using PSD-bin and bandpower features.

2.5.1 TFSNet Architecture and Implementation Details

The TFSNet architecture consists of two parallel feature-processing branches: the Power Spectral Density (PSD) branch and the bandpower branch. The input comprises 640 PSD features (16 EEG channels $\times$ 40 frequency bins) and 80 bandpower features (16 EEG channels $\times$ 5 frequency bands). Each branch employs two 2D convolutional layers with kernel sizes of [specify kernel size for each layer], followed by Batch Normalization (BN) and Rectified Linear Unit (ReLU) activation. A Convolutional Block Attention Module (CBAM) is then applied to enhance relevant channel-wise and spatial features. Adaptive Average Pooling is used to obtain fixed-size feature representations, which are subsequently flattened or tokenized and projected into a 64-dimensional embedding space. Each branch contains one Transformer Encoder layer to capture relationships among the extracted features. The resulting PSD and bandpower representations are integrated through Adaptive

Cross-Domain Attention (ACDA), producing a fused 64-dimensional feature vector. Finally, a Multilayer Perceptron (MLP) consisting of a 64-unit hidden layer with ReLU activation and a single-unit output layer with sigmoid activation generates the probability of ASD classification.

2.5.2 Training Configuration and Model Selection

TFSNet training was conducted independently for each cross-validation fold, with network parameters reinitialized for each fold and a random seed of 42 used to ensure reproducibility. The model was trained using the Adam optimizer with a learning rate of 0.0005, a batch size of 32, and 50 epochs, using the cross-entropy loss function. The best model was selected based on the lowest validation loss, with validation accuracy also monitored to evaluate classification performance. The same training configuration was applied consistently to both denoising approaches, MSPCA and MSICA, to ensure a fair comparison of classification performance.

2.6 Statistical Analysis

Statistical analysis was conducted to determine whether the classification performance difference between TFSNet with MSPCA and TFSNet with MSICA was statistically significant. The analysis was performed using the accuracy values obtained from the 8-fold subject-wise cross-validation. In this study, each fold produced one accuracy value for TFSNet with MSPCA and one corresponding accuracy value for TFSNet with MSICA. Therefore, the two sets of accuracy values were treated as paired observations [27, 28].

The paired difference between the two denoising methods for each fold was calculated as follows:

$$d_i = P_{i,\mathrm{MSPCA}} - P_{i,\mathrm{MSICA}} \tag{15}$$

where $\mathrm{Acc}_{(i,\mathrm{MSPCA})}$ and $\mathrm{Acc}_{(i,\mathrm{MSICA})}$ represent the classification accuracy obtained in the i-th fold using MSPCA and MSICA, respectively. The paired t-test was applied to evaluate whether the mean difference between the two paired accuracy values was statistically significant. The paired t-test statistic is defined as:

$$t = \frac{d}{s_d/\sqrt{n}} \tag{16}$$

where d denotes the mean of the paired differences, s_d represents the standard deviation of the paired differences, and n is the number of folds. In this study, $n = 8$, corresponding to the 8-fold subject-wise cross-validation protocol. In addition to the paired t-test, the Wilcoxon signed-rank test was also used as a non-parametric statistical test. This test was applied to validate the significance of the performance difference without assuming that the paired differences followed a normal distribution. Both statistical tests were evaluated using a significance level of $\alpha = 0.05$. A p-value lower than 0.05 indicates that the performance difference between TFSNet with MSPCA and TFSNet with MSICA is statistically significant.

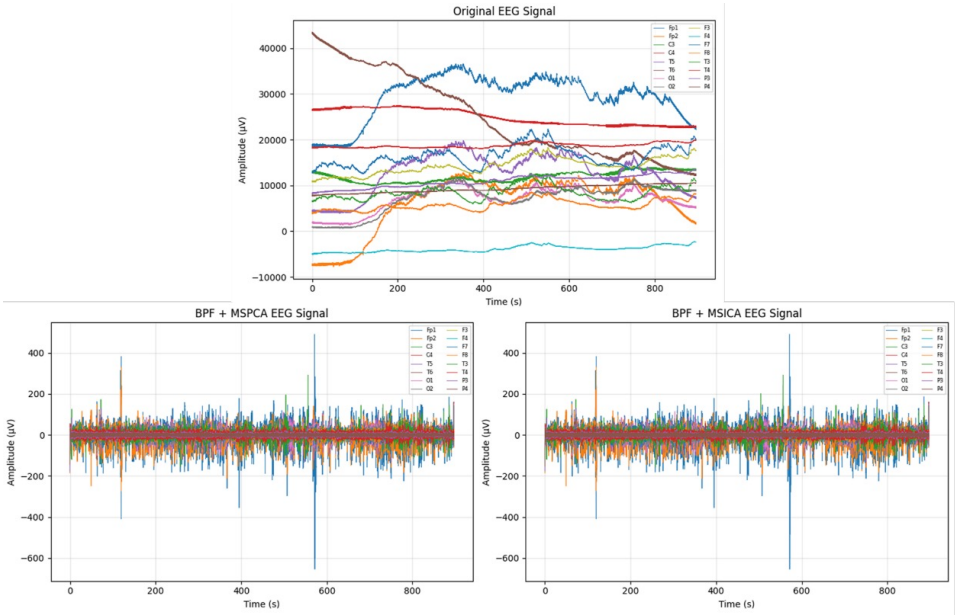

Figure 4: Visualization of EEG signal preprocessing results using BPF + MSPCA and BPF + MSICA.

3 Results

3.1 Preprocessing

Figure 4 shows the preprocessing results of the 16-channel EEG signals using BPF + MSPCA and BPF + MSICA. The original EEG signal exhibits large-amplitude variations and unstable baseline trends, with several channels exceeding 40,000 µV, indicating substantial drift and noise.

Following the application of BPF + MSPCA, the EEG signal exhibited a more stable amplitude centered around 0 μV, indicating a reduction in baseline drift. Similar results were obtained with BPF + MSICA, showing more controlled amplitude fluctuations compared to the original signal. Signal-to-Noise Ratio (SNR) evaluation revealed that BPF + MSICA yielded an average value of 24.4843 dB, which was higher than the 22.2433 dB achieved by BPF + MSPCA. However, this higher SNR did not translate into improved classification performance. TFSNet with MSPCA achieved an accuracy of 98.60%, surpassing the 98.07% accuracy achieved by TFSNet with MSICA. These findings indicate that a higher SNR does not necessarily result in superior classification accuracy. It is hypothesized that MSPCA is more effective at preserving the EEG characteristics relevant for distinguishing between ASD and neurotypical patterns, whereas MSICA may attenuate certain signal components that facilitate the classification process [15]. The 98.60% accuracy achieved by TFSNet with MSPCA demonstrates competitive performance relative to previous EEG-based ASD classification approaches, although direct comparisons must account for differences in dataset characteristics, preprocessing methods, and evaluation protocols [7, 8, 11, 15].

3.2 Classification

3.2.1 TFSNet With MSPCA

The classification performance of TFSNet with Multiscale Principal Component Analysis (MSPCA) denoising is presented in Table 1. Before the classification stage, the EEG signals processed using MSPCA achieved a Signal-to-Noise Ratio (SNR) value of 22.2433 dB. This result indicates that MSPCA improved EEG signal quality by reducing noise while preserving relevant signal information for ASD classification.

Table 1: Classification performance of TFSNet with MSPCA using 8-fold subject-wise cross-validation

Fold	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Specificity (%)	AUC (%)
1	97.93	98.97	96.88	97.91	98.99	99.82
2	98.94	98.45	99.44	98.95	98.44	99.95
3	99.16	99.22	99.11	99.16	99.22	99.96
4	98.83	98.66	98.99	98.83	98.66	99.84
5	98.66	98.77	98.55	98.66	98.77	99.85
6	98.83	98.77	98.88	98.83	98.77	99.88
7	98.10	98.98	97.21	98.08	98.99	99.70
8	98.32	98.43	98.21	98.32	98.44	99.83
Mean	**98.60**	**98.78**	**98.41**	**98.59**	**98.79**	**99.85**

Based on the 8-fold subject-wise cross-validation results, TFSNet with MSPCA achieved consistently high classification performance across all folds. The average accuracy, precision, recall, F1-score, specificity, and AUC were 98.60%, 98.78%, 98.41%, 98.59%, 98.79%, and 99.85%, respectively. The highest performance was obtained in Fold 3, with an accuracy of 99.16%, precision of 99.22%, recall of 99.11%, F1-score of 99.16%, specificity of 99.22%, and AUC of 99.96%. Meanwhile, the lowest performance was observed in Fold 1, with an accuracy of 97.93%, precision of 98.97%, recall of 96.88%, F1-score of 97.91%, specificity of 98.99%, and AUC of 99.82%. Although Fold 1 produced the lowest accuracy, the performance remained above 97%, indicating that TFSNet with MSPCA maintained stable classification capability across different data partitions.

The confusion matrix results for the lowest- and highest-performing folds of TFSNet with MSPCA are shown in Figure 5. In Fold 1, the lowest-performing fold, the model correctly classified 886 normal samples and 868 ASD samples. However, 9 normal samples were misclassified as ASD, while 28 ASD samples were misclassified as normal. This indicates that the dominant error in Fold 1 was false negatives, consistent with the lower recall in this fold.

In contrast, Fold 3 shows the best classification result for TFSNet with MSPCA. The model correctly classified 888 normal samples and 887 ASD samples, with only 7 false positives and 8 false negatives. The small number of misclassified samples in this fold confirms that MSPCA-denoised EEG signals provided highly discriminative information for TFSNet. Overall, the high accuracy, balanced precision-recall values, and strong AUC indicate that TFSNet with MSPCA is effective in distinguishing ASD and normal EEG patterns.

3.2.2 TFSNet with MSICA

The classification performance of TFSNet with Multiscale Independent Component Analysis (MSICA) denoising is summarized in Table 2. Prior to classification, the MSICA-denoised EEG signals achieved an SNR value of 24.4843 dB. This SNR value was higher than that obtained using MSPCA, indicating that MSICA provided stronger noise suppression and improved EEG signal quality before the classification stage.

Table 2: Classification performance of TFSNet with MSICA using 8-fold subject-wise cross-validation

Fold	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Specificity (%)	AUC (%)
1	97.38	98.63	96.09	97.34	98.66	99.75
2	97.77	97.98	97.54	97.76	97.99	99.80
3	98.32	98.32	98.32	98.32	98.32	99.89
4	98.77	98.99	98.55	98.77	98.99	99.85
5	98.16	98.00	98.32	98.16	97.99	99.88
6	97.99	98.20	97.77	97.98	98.21	99.76
7	97.93	97.77	98.10	97.94	97.77	99.82
8	98.21	98.32	98.10	98.21	98.32	99.72
Mean	**98.07**	**98.28**	**97.85**	**98.06**	**98.03**	**99.81**

The 8-fold subject-wise cross-validation results show that TFSNet with MSICA also achieved strong and stable classification performance. The model obtained an average accuracy of 98.07%, precision of 98.28%, recall of 97.85%, F1-score of 98.06%, specificity of 98.03%, and AUC of 99.81%. The highest performance was achieved in Fold 4, with an accuracy of 98.77%, precision of 98.99%, recall of 98.55%, F1-score of 98.77%, specificity of 98.99%, and AUC of 99.85%. Meanwhile, the lowest performance was obtained in Fold 1, with an accuracy of 97.38%, precision of 98.63%, recall of 96.09%, F1-score of 97.34%, specificity of 98.66%, and AUC of 99.75%. These results indicate that TFSNet with MSICA maintained high classification performance across all folds, although the average values were slightly lower than those obtained using MSPCA.

The confusion matrix results for TFSNet with MSICA are shown in Figure 5. In Fold 1, the lowest-performing fold, the model correctly classified 883 normal and 861 ASD samples. In this fold, 12 normal samples were misclassified as ASD, while 35 ASD samples were misclassified as normal. The higher number of false negatives explains the lower recall in Fold 1 and indicates that more ASD samples were misclassified as normal than in the best-performing fold.

In Fold 4, which corresponds to the highest-performing fold for TFSNet with MSICA, the model correctly classified 886 normal and 882 ASD samples. Only 9 normal samples were incorrectly classified as ASD, and 13 ASD samples were incorrectly classified as normal. Compared with Fold 1, Fold 4 produced fewer classification errors and showed a more balanced distribution between false positives and false negatives. These results confirm that MSICA-denoised EEG signals can also provide reliable input for TFSNet in ASD classification. However, despite its higher SNR, the average classification performance of TFSNet with MSICA was slightly lower than that with MSPCA, suggesting that higher

denoising quality, as measured by SNR, does not always directly correspond to higher classification accuracy.

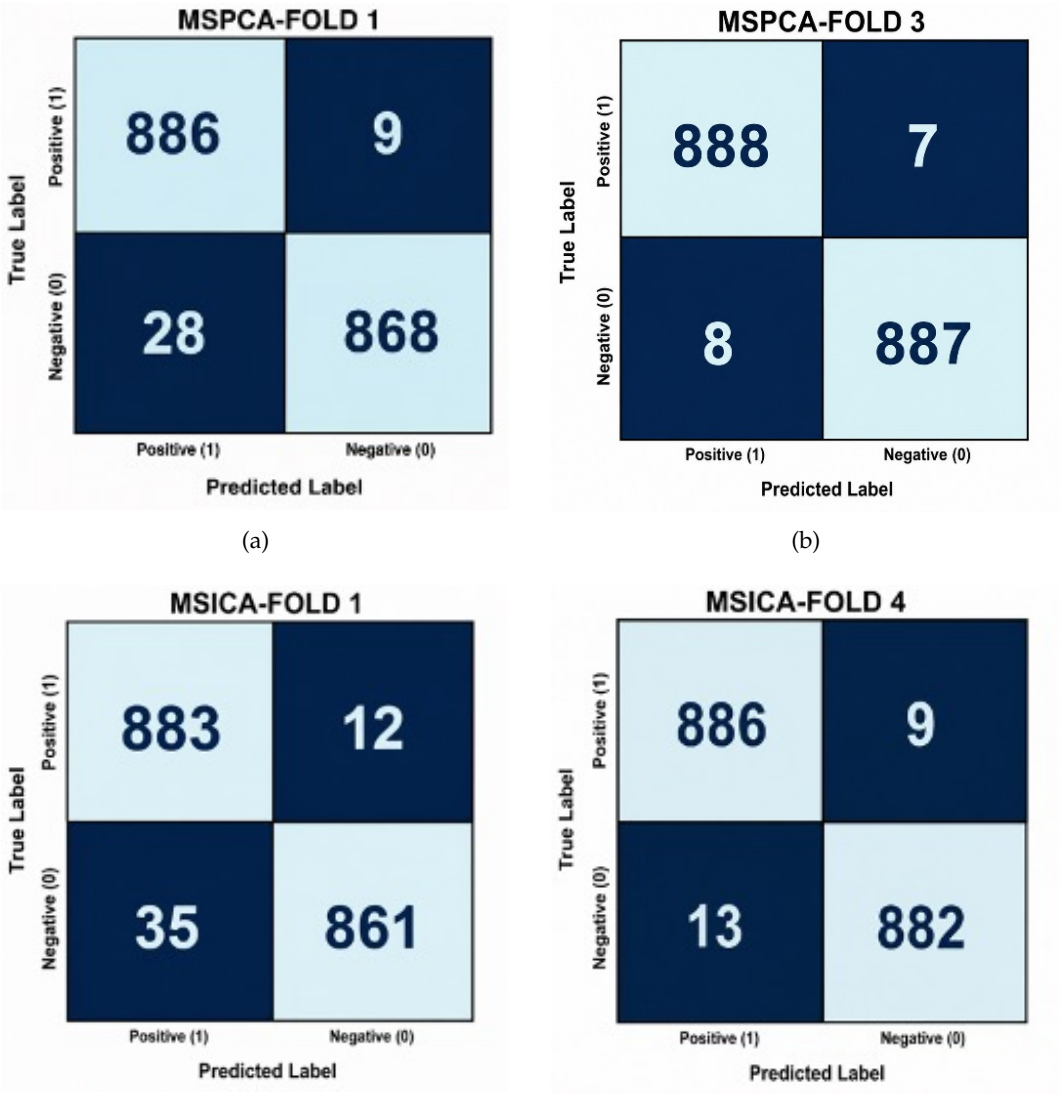

Figure 5: Confusion matrices of TFSNet classification for the selected lowest- and highest-performing folds: (a) TFSNet with MSPCA in fold 1, (b) TFSNet with MSPCA in fold 3, (c) TFSNet with MSICA in fold 1, and (d) TFSNet with MSICA in fold 4.

3.2.3 Statistical Test

To further evaluate whether the performance difference between TFSNet with MSPCA and TFSNet with MSICA was statistically significant, paired t-tests and Wilcoxon signed-rank tests were conducted. The statistical test was performed based on the fold-wise, subject-wise accuracy values obtained from the 8-fold subject-wise cross-validation. The paired t-test was used to examine the mean difference between the two paired classification results, while the Wilcoxon signed-rank test was used as a nonparametric test to assess the significance of the performance difference.

Table 3: Statistical test results comparing TFSNet with MSPCA and MSICA based on accuracy

Model	Paired t-test	P-value	Wilcoxon signed rank (Interpretation)
TFSNet	3.7310	0.0073	P = 0.0078 (Significant)

As shown in Table 3, the paired t-test produced a t-value of 3.7310 with a p-value of 0.0073. In addition, the Wilcoxon signed-rank test yielded a p-value of 0.0078. Since both p-values were below the 0.05 significance threshold, the difference between TFSNet with MSPCA and TFSNet with MSICA was statistically significant. These results indicate that using MSPCA as the denoising method provided a significant improvement in classification accuracy compared with MSICA under the 8-fold subject-wise cross-validation scheme.

Although MSICA achieved a higher SNR than MSPCA, the statistical test shows that MSPCA yielded significantly better classification performance when combined with TFSNet. This finding suggests that the denoising method with the highest SNR does not always yield the highest classification accuracy. In this study, MSPCA may have preserved more discriminative EEG components that are more suitable for TFSNet in distinguishing ASD from normal EEG patterns.

4 Discussion

4.1 Performance of TFSNet with MSPCA and MSICA

This study evaluated the performance of TFSNet for EEG-based classification of Autism Spectrum Disorder (ASD) using two denoising methods: Multiscale Principal Component Analysis (MSPCA) and Multiscale Independent Component Analysis (MSICA). Both denoising methods were applied to improve the quality of EEG signals before classification. The denoising evaluation showed that MSPCA achieved an SNR of 22.2433 dB, while MSICA achieved an SNR of 24.4843 dB. This result indicates that MSICA provided stronger noise suppression, as measured by signal quality. However, the classification results showed that TFSNet with MSPCA achieved slightly higher performance than TFSNet with MSICA across most evaluation metrics.

Based on the 8-fold subject-wise cross-validation, TFSNet with MSPCA achieved an average epoch-level accuracy of 98.60%, with a precision of 98.78%, recall of 98.41%, F1-score of 98.59%, specificity of 98.79%, and AUC of 99.85%. Meanwhile, TFSNet with MSICA achieved an average epoch-level accuracy of 98.07%, precision of 98.28%, recall of 97.85%,

F1-score of 98.06%, specificity of 98.03%, and AUC of 99.81%. These results indicate that both denoising methods provided highly reliable input signals for TFSNet, as all average performance metrics remained above 97%. Nevertheless, MSPCA produced a slightly better classification performance, particularly in accuracy, recall, F1-score, and specificity.

The higher performance of TFSNet with MSPCA suggests that the MSPCA-denoised EEG signals may preserve more discriminative components required by the model to differentiate between ASD and normal EEG patterns. In contrast, although MSICA produced a higher SNR, its classification performance was slightly lower. This finding shows that the highest signal quality, as measured by SNR, does not always correspond to the highest classification accuracy. In EEG-based deep learning classification, preserving class-relevant information is as important as removing noise. Therefore, a denoising method should not only suppress artifacts but also preserve the temporal and spatial characteristics useful for classification.

4.2 Confusion Matrix Analysis

The confusion matrix analysis provides further insight into the classification behavior of TFSNet with MSPCA and MSICA. For TFSNet with MSPCA, the lowest-performing fold was Fold 1, where the model correctly classified 886 normal samples and 868 ASD samples, while 9 normal samples were misclassified as ASD and 28 ASD samples were misclassified as normal. This indicates that the main error in Fold 1 was due to false-negative predictions, consistent with its lower recall. Meanwhile, the highest-performing fold was Fold 3, where the model correctly classified 888 normal samples and 887 ASD samples, with only 7 false positives and 8 false negatives. This result demonstrates that the classification errors were very low and relatively balanced between the two classes.

For TFSNet with MSICA, the lowest-performing fold was also Fold 1, where 883 normal samples and 861 ASD samples were correctly classified. In this fold, 12 normal samples were misclassified as ASD, and 35 ASD samples were misclassified as normal. The higher number of false negatives explains the lower recall value obtained in this fold. In contrast, the highest-performing fold for MSICA was Fold 4, where the model correctly classified 886 normal samples and 882 ASD samples, with only 9 false positives and 13 false negatives. Compared with the lowest-performing fold, Fold 4 had fewer classification errors and a more balanced distribution of false positives and false negatives.

Overall, the confusion matrix results support the quantitative performance metrics obtained from the 8-fold subject-wise cross-validation. The dominant diagonal values in the confusion matrices indicate that TFSNet correctly classified a high proportion of EEG epochs from both ASD and neurotypical participants within the evaluated cohort. Moreover, the relatively small number of false positives and false negatives suggests that the model did not strongly favor one class over the other. This balanced epoch-level classification behavior is relevant for assessing the discriminative capability of the proposed model; however, it should not be interpreted as subject-level diagnostic performance. From an application perspective, false-negative predictions represent misclassified ASD eoochs, whereas false-positive predictions represent neurotypical epochs incorrectly classified as ASD.

4.3 Statistical Significance of Classification Performance

To evaluate whether the performance difference between TFSNet with MSPCA and TFSNet with MSICA was statistically significant, a paired t-test and Wilcoxon signed-rank test were conducted based on the fold subject-wise accuracy values. The paired t-test produced a t-value of 3.7310 with a p-value of 0.0073. In addition, the Wilcoxon signed-rank test yielded a p-value of 0.0078. Since both p-values were below the 0.05 significance threshold, the difference between the two denoising approaches was statistically significant.

Statistical comparisons were performed using the paired accuracy values obtained from the eight subject-wise cross-validation folds. For each fold, the accuracy of TFSNet with MSPCA was paired with the corresponding accuracy of TFSNet with MSICA because both approaches were evaluated on the same held-out subjects. The paired t-test was used to assess the mean difference in accuracy, while the Wilcoxon signed-rank test was employed as a complementary non-parametric analysis. However, the eight folds should not be assumed to be fully independent statistical observations because their training sets overlap. Therefore, the resulting p-values should be interpreted cautiously, particularly given the limited number of folds and participants. The statistical comparison provides preliminary evidence of differences between the two denoising approaches within the evaluated dataset, rather than definitive evidence of generalizable superiority. Further validation using larger, independent cohorts is needed to establish the robustness of this comparison.

Table 4: Statistical comparison of TFSNet with MSPCA and MSICA

Comparison	**Mean Difference**	**95% CI**	**Cohen's d_z**	**Paired t**	**p-value**	**Wilcoxon p**
MSPCA vs MSICA	0.53	0.19–0.87	1.32	3.731	0.0073	0.0078

Table 4 summarizes the statistical comparison between the two preprocessing approaches. The mean fold-wise accuracy difference between TFSNet with MSPCA and MSICA was 0.53 percentage points, with a 95% confidence interval of 0.19–0.87 percentage points. The corresponding paired effect size was large (Cohen's d_z = 1.32). The paired t-test yielded $t(7)$ = 3.731, p = 0.0073, while the Wilcoxon signed-rank test yielded p = 0.0078. These results indicate that the observed fold-wise performance advantage of MSPCA was consistent across the eight evaluation folds. However, because the cross-validation folds share training subjects, the eight fold-wise observations should not be regarded as fully independent samples. Therefore, the statistical results are interpreted as fold-wise comparative evidence of the observed performance difference rather than definitive population-level statistical inference.

4.4 Implications of the Proposed TFSNet-Based Classification

The findings of this study demonstrate that the combination of multiscale denoising and TFSNet holds significant potential for EEG-based ASD classification. EEG signals are highly non-stationary and susceptible to artifacts, making pre-processing a crucial step prior to classification. Applying MSPCA or MSICA before TFSNet enhances signal quality, enabling the model to learn more relevant temporal and spatial EEG patterns. The high values for accuracy, precision, recall, specificity, F1-score, and AUC achieved in this study

indicate that the proposed approach effectively distinguishes between ASD and normal EEG signals.

From a practical perspective, the high classification performance indicates the potential of the proposed TFSNet framework for further investigation in EEG-based ASD classification. However, these results should not be interpreted as evidence of clinical readiness, as the study included only 16 participants from a single cohort. Although subject-wise cross-validation was employed to reduce subject-dependent data leakage, the evaluation does not establish generalizability to independent populations or real-world clinical settings. Furthermore, the reported metrics represent epoch-level classification performance rather than subject-level diagnostic accuracy. Therefore, the proposed framework should be considered a preliminary computational approach. External validation using independent datasets, larger and more diverse cohorts, and participant-level evaluation is necessary before its clinical applicability can be determined.

The integration of multiscale EEG denoising and TFSNet at the feature level demonstrates promising performance in ASD classification, achieving 98.60% accuracy using the MSPCA method and subject-based eight-fold cross-validation. However, these findings should be interpreted with caution given the limitations of the single-cohort dataset, which included only 16 participants. Although subject-based validation can mitigate the risk of data leakage, it does not guarantee generalizability to independent populations or real-world clinical settings. Furthermore, classification accuracy at the segment level does not necessarily reflect diagnostic performance at the patient level. Therefore, the proposed framework should be viewed as a preliminary computational approach rather than a clinically validated diagnostic tool. External validation using larger, more diverse datasets and evaluation at the participant level is required before clinical implementation can be established.

4.5 Limitation and Future Work

A major limitation of this study is the small sample size, consisting of only 16 children, including 8 children with ASD and 8 typically developing controls, recruited from a single cohort. Although the proposed TFSNet achieved an average epoch-level accuracy of 98.60% using 8-fold subject-wise cross-validation, this result should be interpreted as preliminary evidence of classification performance within the collected dataset rather than evidence of clinical applicability. The use of subject-wise cross-validation reduces the risk of data leakage between training and testing subjects but does not guarantee generalization to independent populations. In addition, the large number of EEG epochs generated from each participant does not increase the number of independent participants. Future studies should therefore evaluate the proposed framework using independent external datasets and larger, more diverse, preferably multicenter cohorts. Participant-level evaluation should also be conducted to determine whether epoch-level predictions can provide reliable subject-level classification.

Another limitation is that the reported classification performance is based on EEG epoch-level evaluation. Although subject-wise cross-validation was used to prevent EEG epochs from the same participant from appearing in both training and testing sets, multiple epochs from the same participant may still exhibit within-subject dependence. Therefore, the reported accuracy, precision, recall, F1-score, specificity, and AUC should be interpreted as epoch-level performance rather than subject-level diagnostic performance. Future stud-

ies should investigate subject-level aggregation strategies to determine whether the observed epoch-level performance translates into reliable participant-level classification.

In addition, the present study evaluates two multiscale denoising approaches, namely MSPCA and MSICA. Future research may include additional denoising and preprocessing techniques to provide a broader comparison of their effects on EEG classification. The interpretability of the proposed TFSNet also remains limited, particularly regarding the specific EEG channels, frequency characteristics, and temporal patterns that contribute to the model's decisions. Explainable artificial intelligence, channel contribution analysis, and attention visualization may therefore be investigated in future studies. Finally, the current evaluation was conducted using a single cohort acquired under a specific recording setting. Therefore, external validation using independent datasets and larger multi-center cohorts is necessary to assess the robustness and generalizability of the proposed framework. The current findings should consequently be regarded as preliminary evidence of the feasibility of the proposed approach rather than as evidence of clinical readiness.

5 Conclusion

This study demonstrates that integrating the Time-Frequency Synergy Network (TFSNet) with multiscale denoising methods enables the high-performance classification of ASD and neurotypical EEG patterns. Although MSICA yielded a higher Signal-to-Noise Ratio (SNR) (24.48 dB) compared to MSPCA (22.24 dB), TFSNet combined with MSPCA achieved superior epoch-level classification accuracy: 98.60% versus 98.07% for MSICA. This difference was statistically significant according to both the paired t-test ($p = 0.0073$) and the Wilcoxon test ($p = 0.0078$). These findings indicate that an increase in SNR does not necessarily translate to improved classification accuracy; therefore, the effectiveness of denoising methods should be evaluated based on their ability to support classification performance.

Despite the promising classification results, several limitations must be acknowledged. The evaluation was conducted on a small, single-cohort dataset comprising 16 children, and the reported metrics represent epoch-level rather than subject-level classification performance. Consequently, the results do not establish clinical diagnostic validity or generalizability to independent populations. Future work should prioritize external validation with independent datasets, larger, more diverse cohorts, and participant-level evaluation before the proposed framework can be considered for clinical applications

Acknowledgments

This research was supported by the Institute for Research and Community Service (LPPM), Universitas Syiah Kuala. The authors greatly appreciate the institutional support provided.

References

[1] J. Li, X. Kong, L. Sun, X. Chen, G. Ouyang, X. Li, and S. Chen, "Identification of autism spectrum disorder based on electroencephalography: A systematic review," *Computers in Biology and Medicine*, vol. 170, p. 108075, 2024.

[2] S. Y. Ke, H. Wu, H. Sun, A. Zhou, J. Liu, X. Zheng, K. Liu, M. B. Westover, H. Xu, and X.-j. Kong, "Classification of autism spectrum disorder using electroencephalography in chinese children: a cross-sectional retrospective study," *Frontiers in Neuroscience*, vol. 18, p. 1330556, 2024.

[3] J. Rogala, J. Żygierewicz, U. Malinowska, H. Cygan, E. Stawicka, A. Kobus, and B. Vanrumste, "Enhancing autism spectrum disorder classification in children through the integration of traditional statistics and classical machine learning techniques in eeg analysis," *Scientific Reports*, vol. 13, no. 1, p. 21748, 2023.

[4] A. Lakhan, M. A. Mohammed, K. H. Abdulkareem, H. Hamouda, and S. Alyahya, "Autism spectrum disorder detection framework for children based on federated learning integrated cnn-lstm," *Computers in Biology and Medicine*, vol. 166, p. 107539, 2023.

[5] M. N. A. Tawhid, S. Siuly, H. Wang, F. Whittaker, K. Wang, and Y. Zhang, "A spectrogram image based intelligent technique for automatic detection of autism spectrum disorder from eeg," *Plos one*, vol. 16, no. 6, p. e0253094, 2021.

[6] M. Baygin, S. Dogan, T. Tuncer, P. D. Barua, O. Faust, N. Arunkumar, E. W. Abdulhay, E. E. Palmer, and U. R. Acharya, "Automated asd detection using hybrid deep lightweight features extracted from eeg signals," *Computers in Biology and Medicine*, vol. 134, p. 104548, 2021.

[7] L. Shi, L. Ma, J. Zhao, Z. Kuang, S. Wang, H. Yang, H. Wang, Q. Han, and L. Sun, "Tfsnet: A time–frequency synergy network based on eeg signals for autism spectrum disorder classification," *Brain Sciences*, vol. 15, no. 7, p. 684, 2025.

[8] W. Hu, G. Jiang, J. Han, X. Li, and P. Xie, "Regional-asymmetric adaptive graph convolutional neural network for diagnosis of autism in children with resting-state eeg," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 200–211, 2023.

[9] R. S. Lalawat and V. Bajaj, "An automatic framework for detecting autism spectrum disorder from eeg signals using tfd," *IEEE Sensors Journal*, vol. 24, no. 7, pp. 10632–10639, 2024.

[10] T. Tang, C. Li, S. Zhang, Z. Chen, L. Yang, Y. Mu, J. Chen, P. Xu, D. Gao, F. Li, *et al.*, "A hybrid graph network model for asd diagnosis based on resting-state eeg signals," *Brain Research Bulletin*, vol. 206, p. 110826, 2024.

[11] Y. Xu, Z. Yu, Y. Li, Y. Liu, Y. Li, and Y. Wang, "Autism spectrum disorder diagnosis with eeg signals using time series maps of brain functional connectivity and a combined cnn–lstm model," *Computer methods and programs in biomedicine*, vol. 250, p. 108196, 2024.

[12] S. K. Veeramalla, V. D. R. Tatiparthi, E. B. Babu, R. Sahoo, and T. H. Rao, "Artifact removal of eeg data using wavelet total variation denoising and independent component analysis," *Analog Integrated Circuits and Signal Processing*, vol. 122, no. 2, p. 14, 2025.

[13] M. Klug, T. Berg, and K. Gramann, "Optimizing eeg ica decomposition with data cleaning in stationary and mobile experiments," *Scientific Reports*, vol. 14, no. 1, p. 14119, 2024.

[14] M. M. Rahmat, Y. Nurdin, M. Melinda, Y. Away, M. Irhamsyah, W. Wong, *et al.*, "Autism eeg signal pre-processing: Performance evaluation of ms-ica and butterworth filter," *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 3, pp. 488–500, 2025.

[15] M. M. Rahmat, Y. Yunidar, and M. Melinda, "A comparative study of mspca and msica denoising for eeg based autism spectrum disorder classification using lightgbm and tcn-sa.," *International Journal of Intelligent Engineering & Systems*, vol. 19, no. 6, p. 186, 2026.

[16] M. Melinda, Y. Away, P. D. Purnamasari, F. Fahmi, I. F. Rahman, F. H. Juwono, N. Li, R. Yan, and S. Gazali, "Feature-based eeg classification: A comparative study of emd and ewt with eegnet and shallow fbcsp convnet," *IEEE Access*, 2025.

[17] M. Melinda, P. D. Purnamasari, F. Fahmi, E. P. Sinulingga, M. Muliyadi, Y. Away, Y. Yunidar, and F. H. Juwono, "A comprehensive eeg dataset and performance assessment for autism spectrum disorder: M. melinda et al.," *Scientific reports*, vol. 15, no. 1, p. 34981, 2025.

[18] N. S. Amer and S. B. Belhaouari, "Eeg signal processing for medical diagnosis, healthcare, and monitoring: A comprehensive review," *IEEE Access*, vol. 11, pp. 143116–143142, 2023.

[19] S. Peketi and S. B. Dhok, "Machine learning enabled p300 classifier for autism spectrum disorder using adaptive signal decomposition," *Brain Sciences*, vol. 13, no. 2, p. 315, 2023.

[20] M. Zangeneh Soroush, P. Tahvilian, M. H. Nasirpour, K. Maghooli, K. Sadeghniiat-Haghighi, S. Vahid Harandi, Z. Abdollahi, A. Ghazizadeh, and N. Jafarnia Dabanloo, "Eeg artifact removal using sub-space decomposition, nonlinear dynamics, stationary wavelet transform and machine learning algorithms," *Frontiers in Physiology*, vol. 13, p. 910368, 2022.

[21] G. Zhang, C. D. Carrasco, K. Winsler, B. Bahle, F. Cong, and S. J. Luck, "Assessing the effectiveness of spatial pca on svm-based decoding of eeg data," *NeuroImage*, vol. 293, p. 120625, 2024.

[22] W. S. Neo, D. Foti, B. Keehn, and B. Kelleher, "Resting-state eeg power differences in autism spectrum disorder: a systematic review and meta-analysis," *Translational Psychiatry*, vol. 13, no. 1, p. 389, 2023.

[23] R. Sharma and H. K. Meena, "Emerging trends in eeg signal processing: A systematic review," *SN Computer Science*, vol. 5, no. 4, p. 415, 2024.

[24] X. Geng, X. Fan, Y. Zhong, M. F. Casanova, E. M. Sokhadze, X. Li, and J. Kang, "Abnormalities of eeg functional connectivity and effective connectivity in children with autism spectrum disorder," *Brain Sciences*, vol. 13, no. 1, p. 130, 2023.

[25] Y. Song, Q. Zheng, B. Liu, and X. Gao, "Eeg conformer: Convolutional transformer for eeg decoding and visualization," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 710–719, 2022.

[26] H.-J. Ahn, D.-H. Lee, J.-H. Jeong, and S.-W. Lee, "Multiscale convolutional transformer for eeg classification of mental imagery in different modalities," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 646–656, 2022.

[27] G. Brookshire, J. Kasper, N. M. Blauch, Y. C. Wu, R. Glatt, D. A. Merrill, S. Gerrol, K. J. Yoder, C. Quirk, and C. Lucero, "Data leakage in deep learning studies of translational eeg," *Frontiers in neuroscience*, vol. 18, p. 1373515, 2024.

[28] C. Rommel, J. Paillard, T. Moreau, and A. Gramfort, "Data augmentation for learning predictive models on eeg: a systematic comparison," *Journal of Neural Engineering*, vol. 19, no. 6, p. 066020, 2022.