Fusion of Multi-Band EEG Connectivity Matrix Images for Autism Spectrum Disorder Classification Using Transformer and CNN Models

Zulfan Zulfan [1,*], Junidar Junidar [1], Kikye Martiwi Sukiakhy [1], Aufa Rafiki [2], and Maiza Duana [3]

[1] Department of Informatics, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh, Indonesia
[2] Department of Electrical and Computer Engineering, Faculty of Engineering, Universitas Syiah Kuala, Banda Aceh, Indonesia
[3] Public Health Study Program, Faculty of Health Sciences, Universitas Teuku Umar, Meulaboh, Indonesia
Email: zulfan.abdullah@usk.ac.id (Z.Z.); junidar678@usk.ac.id (J.J.); Kikye.martiwi.sukiakhy@usk.ac.id (K.M.S.); aufa35@mhs.usk.ac.id (A.R.); maizaduana@utu.ac.id (M.D.)
*Corresponding author

***Abstract*—Autism Spectrum Disorder (ASD) is a neurodevelopmental condition whose diagnosis still relies largely on behavioral assessment, creating a need for more objective computer-aided screening approaches. This study proposes a classification framework based on multi-band electroencephalography connectivity matrix images derived from resting-state recordings. Electroencephalography signals from 16 subjects, consisting of 8 subjects with autism spectrum disorder and 8 typically developing controls, were preprocessed using Butterworth bandpass filtering and independent component analysis. The denoised signals were then decomposed into theta, alpha, and beta bands, segmented using overlapping windows, and transformed into phase locking value connectivity matrices. These matrices were converted into connectivity images and classified using two deep learning architectures, namely Data-efficient Image Transformer and ResNet18, under subject-wise 8-fold stratified cross-validation. In the single-band setting, the beta-band images yielded the best performance for both models, achieving mean accuracies of 0.9817 ± 0.0032 for Data-efficient Image Transformer and 0.9689 ± 0.0079 for ResNet18. In the fusion setting, combining theta, alpha, and beta connectivity images further improved performance, producing the highest mean accuracies of 0.9897 ± 0.0048 for Data-efficient Image Transformer and 0.9860 ± 0.0041 for ResNet18. Wilcoxon signed-rank analysis confirmed that the fusion schemes significantly outperformed the best single-band baseline. These findings demonstrate that multi-band electroencephalography connectivity image fusion provides an effective representation for autism spectrum disorder classification and that transformer-based modeling offers a consistent advantage over the convolutional baseline.**

***Keywords*—autism spectrum disorder, electroencephalography, functional connectivity, phase locking value, connectivity matrix image, multi-band fusion, deep learning**

I. Introduction

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition characterized by persistent impairments in social communication and interaction, along with restricted and repetitive patterns of behavior and interests. Because ASD presents substantial heterogeneity across individuals, timely and objective identification remains a major challenge in clinical practice. Early diagnosis is particularly important because it can support earlier intervention and improve developmental outcomes. However, current diagnosis still relies predominantly on behavioral assessment, which is often time-consuming and highly dependent on clinical expertise [1–4].

Among neurophysiological assessment tools, Electroencephalography (EEG) has attracted increasing attention in ASD research because it is non-invasive, relatively affordable, and capable of capturing neural dynamics with high temporal resolution. Recent studies have shown that EEG can reveal atypical oscillatory activity, abnormal synchronization, and altered functional connectivity patterns in individuals with ASD [1, 4–7]. In particular, connectivity-based EEG analysis has emerged as a promising direction because ASD is widely associated with disruptions in large-scale brain communication rather than isolated abnormalities in specific regions alone [5–8]. These findings suggest that EEG-derived connectivity biomarkers may provide clinically meaningful information for computer-aided ASD classification.

Manuscript received April 1, 2026; revised May 9, 2026; accepted July 13, 2026; published November 14, 2026.

In parallel with advances in neurophysiological analysis, machine learning and deep learning approaches have been increasingly applied to ASD identification. Recent studies have explored augmented EEG images, Convolutional Neural Networks (CNNs), graph-based learning, scalogram representations, and connectivity-driven models to improve diagnostic performance [8–11]. Several studies have reported encouraging results by transforming EEG signals into image-like or graph-structured representations, enabling deep models to learn discriminative patterns more effectively [8–10, 12, 13]. This development is particularly relevant to image-oriented research because it shows that the representation of EEG signals can strongly influence classification performance.

Despite this progress, several important gaps remain. First, many previous EEG-based ASD studies have focused on single-band analysis, handcrafted features, or only one family of classification models, which may limit the ability to capture complementary discriminative information distributed across multiple frequency bands [3, 4, 14]. Second, although connectivity analysis has shown considerable promise, its integration with image-based deep learning frameworks is still relatively limited, especially in studies that systematically compare transformer-based vision models with conventional CNN architectures [3, 8, 14]. Third, fusion strategies are increasingly recognized as beneficial in ASD biomarker research, yet fusion at the level of multi-band EEG connectivity images remain insufficiently explored [1, 2, 6]. Consequently, it is still not fully understood how much additional discriminative value can be obtained by combining theta-, alpha-, and beta-band connectivity representations within a unified image-based classification framework.

To address these gaps, this study proposes a fusion-based framework for ASD classification using multi-band EEG connectivity matrix images. Resting-state EEG signals are preprocessed through Butterworth bandpass filtering, independent component analysis, and windowing, followed by decomposition into theta, alpha, and beta frequency bands. For each band, connectivity matrices are generated and represented as images, allowing visual deep learning models to learn band-specific and fused patterns of brain connectivity. Two model families are investigated, namely a transformer-based architecture and a CNN-based architecture, in order to compare their effectiveness on both single-band and fused connectivity images. In this way, the present work aims not only to improve classification performance, but also to clarify the relative contribution of different EEG bands and the added value of multi-band fusion for ASD detection.

The methodological contribution of this study lies in the design and evaluation of a multi-band EEG connectivity image fusion framework, rather than in proposing a new neural network architecture. Specifically, the framework transforms band-specific Phase Locking Value (PLV) connectivity matrices into image representations, aligns theta-, alpha-, and beta-band images at the same subject and window index, and integrates them into a unified three-channel representation for visual deep learning. This representation-level fusion enables complementary frequency-specific connectivity patterns to be learned jointly while preserving subject-wise separation during evaluation.

The main research contributions of this study are summarized as follows. First, this study develops an image-based EEG connectivity representation for ASD classification by transforming theta-, alpha-, and beta-band PLV matrices into connectivity matrix images. Second, it introduces a subject- and window-aligned multi-band fusion strategy that stacks theta, alpha, and beta connectivity images into a unified three-channel representation. Third, it provides a systematic comparison between a transformer-based visual model, namely DeiT, and a CNN-based baseline, namely ResNet18, across both single-band and fused connectivity image schemes. Fourth, it evaluates the proposed framework using subject-wise 8-fold stratified Cross-Validation (CV) with explicit leakage-prevention procedures and subject-level aggregation analysis to examine whether the window-level performance remains consistent at the biological subject level.

II. Materials and Methods

This section describes the overall framework of the proposed study, including the dataset, preprocessing procedures, frequency-band decomposition, connectivity matrix generation, deep learning classification schemes, fusion strategy, and performance evaluation. The complete research pipeline is illustrated in Fig. 1, while the step-by-step methodological procedure is summarized in Algorithm 1.

A. Dataset

This study used an EEG dataset provided by King Abdulaziz University (KAU), Jeddah, Saudi Arabia, consisting of 16 resting-state EEG recordings from 8 participants with Autism Spectrum Disorder (ASD) and 8 typically developing controls [15]. The recordings were acquired using Ag/AgCl electrodes and a g.tec EEG acquisition system under the BCI2000 platform, sampled at 256 Hz, stored in BCI2000 .dat format, and recorded with an acquisition bandpass range of 0.1–60 Hz and a 60 Hz notch filter. Sixteen EEG channels following the international 10–20 system were available and retained in this study, namely Fp1, F3, F7, T3, T5, O1, C4, Fp2, Fz, F4, F8, C3, Cz, Pz, Oz, and O2, covering frontal, temporal, central, parietal, and occipital regions that are relevant to ASD-related spectral and functional connectivity alterations [16–18]. The ASD group consisted of five males and three females aged 6–20 years, with a total recording duration of 4104.2 s, whereas the control group consisted of eight males aged 9–13 years, with a total recording duration of 4534.9 s. The absence of female participants in the control group was inherited from the original KAU dataset and was not an exclusion criterion in the present study. This imbalance may introduce sex-related confounding; therefore, the results should be

interpreted as within-dataset performance under the available subject composition rather than evidence of cross-sex generalizability. Subject-level labels were provided by the clinical team at KAU Hospital; however, the public dataset documentation does not specify the diagnostic instruments used for ASD confirmation, which is acknowledged as a limitation of the source dataset [15].

All recordings were anonymized, and according to the original dataset source, written informed consent was obtained from participants or their legal guardians and the protocol was approved by the KAU Ethics Committee [15]. Since the same dataset has also been used in previous ASD-related EEG classification studies, its use supports comparability with prior work [19, 20].

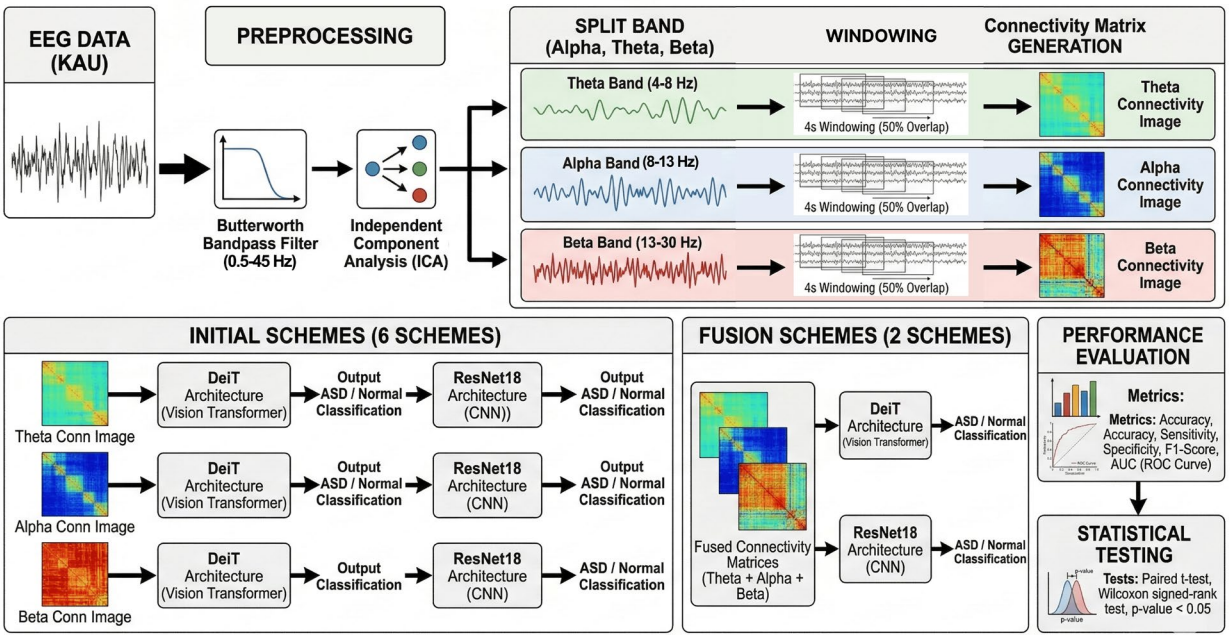


Fig. 1. Overview of the proposed EEG preprocessing and classification framework.

Algorithm 1. Proposed multi-band EEG connectivity image fusion framework

Input: Resting-state EEG recordings with subject IDs and class labels.

Output: Single-band and fused PLV connectivity images, classification results, and evaluation metrics.

1. Apply Butterworth bandpass filtering and Independent Component Analysis (ICA)-based artifact attenuation to each EEG recording.
2. Decompose the cleaned EEG signals into theta, alpha, and beta frequency bands.
3. Segment each band-specific signal into 4 s windows with 50% overlap.
4. For each subject, frequency band, and window, compute pairwise PLV values across all EEG channels.
5. Convert each 16 × 16 PLV matrix into a connectivity image.
6. Assign all window-derived images from the same subject to the same fold under subject-wise 8-fold stratified cross-validation.
7. For the single-band schemes, classify theta, alpha, and beta connectivity images independently.
8. For the fusion schemes, stack theta-, alpha-, and beta-band images from the same subject and window index into a three-channel image.
9. Train and evaluate DeiT and ResNet18 models under the predefined subject-wise folds.
10. Report fold-wise, mean, and subject-level aggregation performance.

Although the dataset contains only 16 subjects, each recording provides continuous multichannel EEG data with sufficient temporal duration to generate multiple fixed-length windows for model training and evaluation. Therefore, the dataset was considered suitable for a controlled proof-of-concept evaluation of the proposed representation and classification framework. However, the biological sample size remains limited; consequently, the results should be interpreted as within-dataset evidence rather than as proof of broad clinical generalizability.

B. Butterworth Bandpass Filtering

Before artifact removal and segmentation, all EEG recordings were filtered using a fourth-order Butterworth bandpass filter to reduce low-frequency drift and high-frequency noise while preserving physiologically relevant EEG components. The cutoff frequencies were set to 0.5 Hz and 45 Hz. The lower cutoff was used to suppress slow baseline wander and electrode drift, whereas the upper cutoff was used to attenuate high-frequency noise and muscle-related interference outside the conventional EEG frequency range commonly analyzed in ASD studies [3, 21, 22]. The Butterworth filter was selected because of its maximally flat passband response, which helps preserve the amplitude characteristics of EEG signals within the frequency range of interest and supports more stable downstream machine learning analysis [21].

To minimize phase distortion, filtering was implemented using a second-order sections representation with forward-backward zero-phase filtering [21]. Let $x[n]$ denote the discrete EEG signal and $H(z)$ the transfer function of the Butterworth bandpass filter. The filtered signal $y[n]$ can be expressed as:

$$y[n] = \mathcal{F}^{-1}\{H(e^{j\omega}) \cdot \mathcal{F}\{x[n]\}\} \quad (1)$$

where $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ represent the discrete Fourier transform and its inverse, respectively. Filtering was applied independently to each EEG channel after removing the first 4 s of each recording to reduce possible acquisition-related transients. The filtered signals were then used as the input for Independent Component Analysis (ICA)-based artifact removal and subsequent segmentation. The same filtering configuration was applied consistently across all experimental settings to ensure fair comparison among preprocessing pipelines and classification schemes.

C. ICA-based Artifact Removal

Following Butterworth bandpass filtering, ICA was applied to reduce residual physiological and non-physiological artifacts, such as eye movements, muscle activity, and high-frequency disturbances, by decomposing multichannel EEG recordings into statistically independent components [21, 23–25]. The main artifacts considered in this study were low-frequency baseline drift, eye-movement or blink-related activity, muscle-related high-frequency activity, and transient non-neural disturbances. Baseline drift and broadband high-frequency noise were first attenuated using the 0.5–45 Hz Butterworth bandpass filter, while residual artifacts were further handled using ICA. Artifact-related components were identified using abnormal kurtosis and excessive high-frequency energy above 25 Hz, because transient movements and muscle-related activity often produce non-Gaussian or high-frequency-dominant components [26, 27]. In this study, ICA was implemented using the FastICA algorithm with unit-variance whitening. The number of independent components was set to the minimum between the number of EEG channels and 16 components, corresponding to the full channel configuration of the dataset. Before decomposition, each EEG channel was standardized to zero mean and unit variance, and FastICA was executed with a fixed random seed and a maximum of 2000 iterations to ensure reproducibility across folds and experimental configurations.

Let the bandpass-filtered EEG signal be represented by:

$$X \in \mathbb{R}^{C \times N} \quad (2)$$

where C is the number of channels and N is the number of temporal samples. ICA models the signal as a linear mixture of statistically independent components:

$$X = AS \quad (3)$$

where $S \in \mathbb{R}^{K \times N}$ contains the independent source components and $A \in \mathbb{R}^{C \times K}$ is the mixing matrix. Artifact-related components were identified using a heuristic scoring procedure based on kurtosis and high-frequency energy ratio, following recent automated and learning-assisted EEG artifact removal approaches [24, 28]. Components with high standardized kurtosis or excessive spectral power above 25 Hz were considered artifact candidates. These components were suppressed by setting their activations to zero, while the remaining components were preserved. The denoised EEG signals were then reconstructed using the original mixing matrix and inverse normalization to restore the original amplitude scale. This ICA-based cleaning was performed after Butterworth filtering and before window segmentation to improve signal quality while maintaining reproducibility and reducing subjective intervention, consistent with recent EEG preprocessing recommendations [23, 25, 27, 29].

To quantify the contribution of the preprocessing stages, an additional ablation experiment was conducted using the alpha-band DeiT setting as a representative configuration. Three input conditions were compared: raw-input condition without global preprocessing, global Butterworth-only condition, and the complete Butterworth + ICA preprocessing pipeline. In the raw-input condition, the global 0.5–45 Hz Butterworth filter and ICA-based artifact attenuation were omitted, while alpha-band decomposition for PLV construction was retained. In the global Butterworth-only condition, the 0.5–45 Hz Butterworth filter was applied without ICA. In the complete condition, both Butterworth filtering and ICA were applied before alpha-band decomposition, windowing, PLV connectivity matrix construction, and image generation. All conditions used the same subject-wise 8-fold cross-validation protocol, fold assignments, windowing strategy, PLV-based connectivity image generation procedure, DeiT architecture, and training configuration, so performance differences could be attributed to the incremental contribution of the preprocessing stages rather than changes in the model, data split, or evaluation protocol. The corresponding ablation results are reported in Section III.A.

D. Frequency Band Decomposition

After global preprocessing, the denoised EEG signals were decomposed into three canonical frequency bands: theta, alpha, and beta. These bands were selected because resting-state EEG activity and connectivity alterations in ASD have frequently been reported in the theta, alpha, and beta ranges, making them relevant for discriminative representation learning [3, 5, 22, 30, 31]. Delta and gamma bands were not included for methodological and signal-quality reasons. The delta band was excluded because the 4 s window used in this study provides relatively few oscillatory cycles at very low frequencies, which may reduce the stability of phase-based PLV estimation and increase sensitivity to residual slow drift, drowsiness, and eye-movement-related activity [30]. The gamma band was excluded because scalp-recorded gamma activity is more susceptible to high-frequency muscle and microsaccadic artifacts, and the present preprocessing pipeline already treated excessive high-frequency power above 25 Hz as a possible artifact-related component [31]. Moreover, because the global preprocessing filter was limited to 0.5–45 Hz, only a narrow lower-gamma range would be available, making gamma-band PLV estimation less robust in this limited-subject dataset. Therefore, theta,

alpha, and beta were selected as a balanced set of frequency bands that are commonly used in resting-state EEG connectivity analysis and are more suitable for stable window-based PLV connectivity image construction [30, 32]. Frequency decomposition was applied after Butterworth filtering and ICA-based artifact attenuation by filtering the cleaned multichannel signals into theta, alpha, and beta sub-bands using a fourth-order Butterworth bandpass filter with zero-phase implementation [3, 21, 22]. The overall band decomposition process is illustrated in Fig. 2.

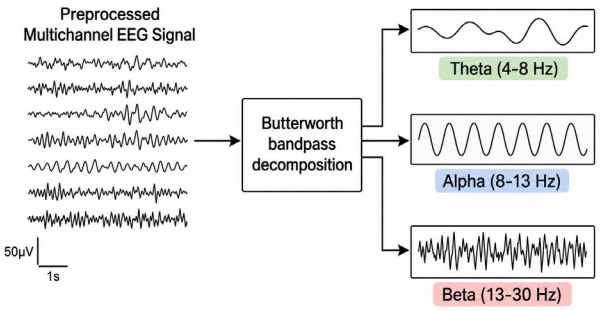


Fig. 2. Schematic illustration of EEG frequency band decomposition.

Let $x_c[n]$ denote the preprocessed EEG signal from channel c, where $c = 1,2,\dots,C$. For each frequency band $b \in \{\theta, \alpha, \beta\}$, the band-limited signal $x_c^{(b)}[n]$ is obtained as:

$$x_c^{(b)}[n] = \mathcal{B}_b\{x_c[n]\} \tag{4}$$

where $\mathcal{B}_b\{\cdot\}$ denotes the bandpass filtering operator associated with the selected frequency range. In this study, the three frequency bands were defined as:

$$\theta: 4 \le f < 8 \text{ Hz} \tag{5}$$

$$\alpha: 8 \le f < 13 \text{ Hz} \tag{6}$$

$$\beta: 13 \le f \le 30 \text{ Hz} \tag{7}$$

Each sub-band signal retained the original multichannel structure while emphasizing rhythm-specific oscillatory activity. This decomposition allowed the subsequent PLV-based connectivity analysis to focus on frequency-specific synchronization patterns rather than broadband activity alone. Such band-specific connectivity representations may reveal complementary information across theta, alpha, and beta rhythms [30, 31, 33]. Following decomposition, each band-limited signal was segmented into 4 s windows with 50% overlap, as described in the next subsection. Thus, for each subject and window, three band-specific multichannel segments were obtained and subsequently used to construct PLV-based functional connectivity matrices. The complete end-to-end workflow is summarized in Fig. 1.

E. Windowing Strategy

After frequency-band decomposition, each 16-channel band-limited EEG signal was segmented into fixed-length overlapping windows for time-series classification. With a sampling frequency of $F_s = 256$ Hz, a 4 s window was used, resulting in $L = 1024$ samples per window, with 50% overlap and a hop size of $H = 512$ samples. This setting was selected to balance temporal coverage, approximate local stationarity, and the need to generate sufficient training samples for deep learning models [34]. The use of overlapping windows also helps reduce boundary effects between adjacent segments and has been reported to improve EEG-based classification performance compared with non-overlapping segmentation [35]. The complete preprocessing and segmentation procedure is illustrated in Fig. 1.

Formally, let $x_c[n]$ denote the preprocessed discrete-time EEG signal from channel c, sampled at $F_s = 256$ Hz. The sample indices of the m-th window are given by:

$$n = mH, mH + 1, \dots, mH + L - 1 \tag{8}$$

and the total number of windows for subject s is computed as:

$$M_s = \left\lfloor \frac{N_s - L}{H} \right\rfloor + 1 \tag{9}$$

where N_s denotes the number of samples in the recording. Each window is represented as a multichannel matrix $X_s^{(m)} \in \mathbb{R}^{C \times L}$, where $C = 16$:

$$X_s^{(m)}(c, i) = x_c[mH + i], c = 1, \dots, 16, i = 0, \dots, L - 1 \tag{10}$$

Windows shorter than 4 s at the end of each recording were discarded to ensure consistent input dimensions. The resulting fixed-size band-specific segments were used to compute PLV-based functional connectivity matrices and then converted into connectivity images. Importantly, each generated window retained its subject ID, and cross-validation was performed at the subject level rather than the window level. Thus, all windows from the same subject were assigned to the same fold, preventing subject-level leakage during evaluation.

F. Subject-wise Cross-validation and Leakage Prevention

To prevent data leakage in this small-sample EEG classification setting, all cross-validation splits were performed strictly at the subject level [3, 36]. Subject ID was used as the grouping variable in an 8-fold stratified cross-validation protocol. Since the dataset contained 8 ASD and 8 Normal subjects, each fold used two held-out test subjects, consisting of one ASD subject and one Normal subject. Thus, each subject appeared in the test partition exactly once.

Although each EEG recording was segmented into overlapping 4 s windows, the windowing process was not used to define the split. All windows, PLV matrices, and connectivity images generated from the same subject were assigned to the same fold, preventing adjacent or highly correlated windows from the same subject from appearing

simultaneously in the training and testing sets. The same subject-level partitioning was applied to both single-band and fusion schemes. In the single-band setting, all theta, alpha, and beta images inherited the fold assignment of their corresponding subject. In the fusion setting, theta-, alpha-, and beta-band images were stacked only when they belonged to the same subject and window index, and the resulting fused image inherited the same subject-level fold assignment. Therefore, no cross-subject or cross-fold mixing occurred, and the held-out test subjects were used only for final fold evaluation.

The adequacy of the data was therefore addressed by using subject-wise rather than window-wise validation. Although windowing increased the number of training instances, these windows were treated as repeated observations from the same biological subject and were never split across training and testing partitions. This protocol allowed the model to learn from multiple temporal segments while ensuring that evaluation was performed on unseen subjects, reducing the risk of subject-level data leakage.

G. PLV Connectivity Matrix Construction and Image Generation

After frequency-band decomposition and band-specific windowing, functional connectivity matrices were constructed for each EEG segment using the Phase Locking Value (PLV). PLV was selected because it measures the consistency of phase synchronization between two EEG time series and has been widely used to characterize rhythmic brain interactions and connectivity-based representations in neurological and neurodevelopmental studies, including ASD-related brain network alterations [5, 7, 13, 31, 33]. In this study, PLV was computed independently for each subject, frequency band, and window. As illustrated in Fig. 3, each band-specific EEG window was transformed into a symmetric 16 × 16 connectivity matrix through Hilbert-based phase extraction, pairwise phase-difference estimation, and PLV computation.

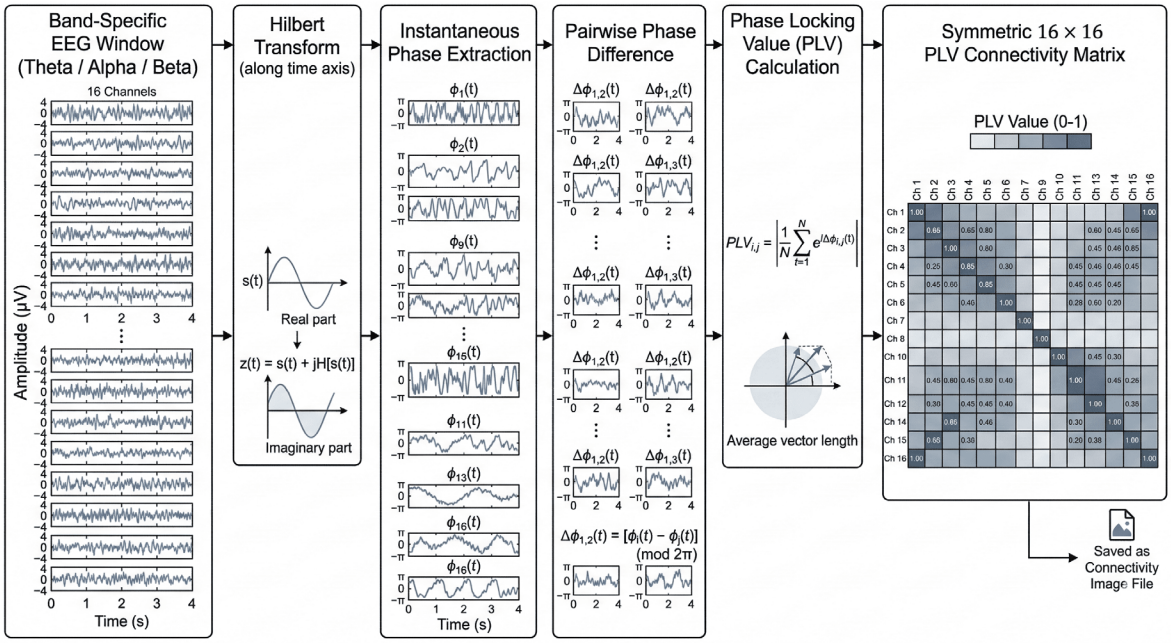


Fig. 3. Schematic illustration of PLV-based connectivity matrix construction from band-specific EEG windows.

Let $X_s^{(m,b)} \in \mathbb{R}^{C\times L}$ denote the m-th window of subject s in frequency band b, where $C = 16$ is the number of EEG channels and L is the number of temporal samples. For each channel c, the analytic signal was obtained using the Hilbert transform:

$$z_c^{(m,b)}[n] = x_c^{(m,b)}[n] + j\,\mathcal{H}\{x_c^{(m,b)}[n]\} \qquad (11)$$

where $x_c^{(m,b)}[n]$ is the real-valued band-limited signal, $\mathcal{H}\{\cdot\}$ denotes the Hilbert transform, and $j = \sqrt{-1}$. The instantaneous phase was then extracted as:

$$\phi_c^{(m,b)}[n] = \angle z_c^{(m,b)}[n] \qquad (12)$$

For each pair of channels p and q, the phase difference sequence was defined as:

$$\Delta\phi_{pq}^{(m,b)}[n] = \phi_p^{(m,b)}[n] - \phi_q^{(m,b)}[n] \qquad (13)$$

The PLV between channels p and q was calculated by averaging the complex phase difference across all samples in the window:

$$\mathrm{PLV}_{pq}^{(m,b)} = \left| \frac{1}{L} \sum_{n=1}^{L} e^{j\Delta\phi_{pq}^{(m,b)}[n]} \right| \qquad (14)$$

The resulting PLV values lie in the interval [0,1], where values close to 1 indicate stable phase synchronization and

values close to 0 indicate weak or inconsistent phase locking. By computing PLV for all channel pairs, a symmetric connectivity matrix $M_s^{(m,b)} \in \mathbb{R}^{C\times C}$ was obtained:

$$M_s^{(m,b)}(p,q) = \mathrm{PLV}_{pq}^{(m,b)}, p,q = 1,2,\ldots,C. \quad (15)$$

The diagonal entries of each matrix were set to 1 to represent self-connectivity. Thus, each window produced one 16 × 16 PLV connectivity matrix for each frequency band. This PLV-based representation was used because it directly captures narrow-band phase synchronization, is computationally suitable for window-based EEG analysis, and has been applied in EEG functional-network construction and ASD-related connectivity studies [10, 13, 34, 35].

PLV was selected as the primary connectivity measure because the objective of this study was to characterize narrow-band phase synchronization between EEG channels. A methodological comparison between PLV and other commonly used EEG connectivity measures is summarized in Table I. Compared with amplitude-based or mixed amplitude-phase measures, PLV focuses directly on phase consistency and is therefore suitable for constructing frequency-specific synchronization matrices [37]. Pearson correlation is simple but mainly captures linear amplitude co-variation, while coherence combines amplitude and phase information and may be influenced by spectral power changes. Phase Lag Index (PLI) and Weighted Phase Lag index (wPLI) are more robust to zero-lag volume conduction effects, but they may discard potentially informative zero-phase synchronization and can be less sensitive in short-window analysis [37, 38]. Therefore, PLV was considered appropriate for the present image-based framework because it provides a simple, bounded, symmetric, and interpretable channel-by-channel matrix representation. However, this study does not claim that PLV is universally superior to other connectivity measures; future work should compare PLV with coherence, PLI, wPLI, correlation-based, and information-theoretic connectivity measures under the same subject-wise validation protocol [39].

TABLE I. COMPARISON OF COMMON EEG CONNECTIVITY MEASURES

Measures	Main Information Captured	Strength	Limitation	Relevance to this Study
PLV	Phase synchronization	Simple, bounded 0–1, symmetric, suitable for phase-based matrix images	Sensitive to zero-lag volume conduction	Selected because this study focuses on narrow-band phase synchrony
Coherence	Amplitude and phase coupling in frequency domain	Captures spectral coupling	Can be affected by amplitude/power variation	Useful alternative, but less phase-specific
Pearson correlation	Linear amplitude co-variation	Simple and computationally efficient	Not phase-specific and sensitive to common-source effects	Less aligned with the phase-synchrony objective
PLI	Non-zero phase-lag synchronization	Reduces zero-lag volume conduction effects	May discard meaningful zero-phase synchronization	Useful robustness comparison
wPLI	Weighted non-zero phase-lag synchronization	More robust to spurious zero-lag coupling than PLV	May require stable phase-lag estimates and can reduce sensitivity to zero-phase effects	Useful future comparison

To construct image-based inputs for deep learning, each PLV matrix was clipped to the interval [0,1] and converted into a heatmap-style connectivity image without axis annotations. Each matrix was exported as a Portable Network Graphics (PNG) file, producing one connectivity image for every subject, window, and frequency band. The generated images were stored according to class label and band type and were later used as inputs for the single-band and fusion classification schemes. As summarized in Fig. 1, the theta-, alpha-, and beta-band PLV matrices were integrated into the overall connectivity image classification workflow.

For reproducibility, the main parameter settings were as follows: EEG sampling frequency was 256 Hz; the global preprocessing bandpass filter was 0.5–45 Hz; the analyzed frequency bands were theta $4 \le f < 8$ Hz, alpha $8 \le f < 13$ Hz, and beta $13 \le f \le 30$ Hz; windowing used 4 s windows with 50% overlap, corresponding to 1024 samples per window and a 512-sample hop size; PLV was computed independently for each subject, frequency band, and window using Hilbert-transform-based phase extraction; pairwise PLV values were computed across 16 EEG channels to form a 16 × 16 matrix with diagonal entries set to 1; and the resulting matrices were clipped to [0,1] and exported as PNG heatmap images without axis annotations.

H. Deep Learning Classification Schemes

To evaluate the discriminative capability of the generated connectivity matrix images, two image-based deep learning architectures were used: Data-efficient Image Transformer (DeiT) and ResNet18. DeiT represents a transformer-based model that can capture long-range dependencies through self-attention, whereas ResNet18 represents a convolutional baseline that learns hierarchical local features through residual convolutional operations. This comparison is relevant because PLV connectivity matrices may contain both global synchronization patterns and localized block-like intensity structures [4, 9, 12, 14]. The use of ImageNet-pretrained weights also provided a transfer-learning strategy suitable for the limited number of EEG-derived connectivity images [4, 9]. ImageNet pretraining was used as a generic visual initialization strategy, not because natural images and PLV connectivity matrices are semantically correlated. The transfer learning contribution in this study lies in the reuse of low- and

mid-level spatial feature initialization, such as edge-like responses, contrast patterns, local textures, and spatial dependency modeling, which can still be useful for structured two-dimensional PLV images. Although PLV matrices represent inter-channel phase synchronization rather than natural objects, they contain organized spatial patterns, including diagonal structure, local block-like textures, intensity gradients, and global inter-channel relationships. Therefore, pretrained visual backbones were used to provide a more stable starting point than random initialization in the limited-data setting. The models were subsequently fine-tuned on the generated PLV connectivity images, allowing the pretrained representations to adapt to EEG-specific connectivity patterns.

As summarized in Fig. 1, two classification settings were evaluated: single-band classification and multi-band fusion classification. In the single-band setting, theta, alpha, and beta connectivity images were classified independently using both DeiT and ResNet18. Let $I_s^{(m,b)}$ denote the connectivity image generated from the m-th window of subject s in frequency band $b \in \{\theta, \alpha, \beta\}$. The predicted label is given by:

$$\hat{y}_s^{(m,b)} = f\left(I_s^{(m,b)}\right) \tag{16}$$

where $\hat{y}_s^{(m,b)} \in \{0,1\}$, with 1 denoting ASD and 0 denoting Normal. The single-band setting therefore consisted of six schemes:

$$f \in \{\text{DeiT}, \text{ResNet18}\}, b \in \{\theta, \alpha, \beta\} \tag{17}$$

For the fusion setting, theta-, alpha-, and beta-band images from the same subject and window index were loaded in grayscale mode and stacked along the channel dimension to form a three-channel representation. Let $I_s^{(m,\theta)}$, $I_s^{(m,\alpha)}$, and $I_s^{(m,\beta)}$ denote the three band-specific images associated with subject s and window m. The fused image is defined as:

$$I_s^{(m,\text{fusion})} = \Phi\left(I_s^{(m,\theta)}, I_s^{(m,\alpha)}, I_s^{(m,\beta)}\right) \tag{18}$$

where $\Phi(\cdot)$ denotes channel-wise stacking. The corresponding prediction is:

$$\hat{y}_s^{(m,\text{fusion})} = f\left(I_s^{(m,\text{fusion})}\right) \tag{19}$$

where $f \in \{\text{DeiT,ResNet18}\}$. This produced two additional schemes: fused-image classification with DeiT and fused-image classification with ResNet18.

All connectivity images were resized from the original 16 × 16 PLV matrices to 224 × 224 pixels to match the input size required by the ImageNet-pretrained DeiT and ResNet18 backbones. This resizing was performed only for architectural compatibility and did not introduce new connectivity information, because the original PLV values and their channel-pair topology remained the source representation. In the single-band setting, PNG images were loaded in Red-Green-Blue (RGB) color format, whereas in the fusion setting, theta, alpha, and beta images were loaded in grayscale and stacked into a three-channel tensor. Inputs were normalized using the standard ImageNet mean and standard deviation values to match the pretrained visual backbones. Geometric augmentation, including horizontal flipping, was not applied because PLV connectivity images have fixed channel-by-channel topology; flipping or rotating them may alter channel-pair relationships and produce physiologically inconsistent patterns.

For the classifier configuration, a DeiT-tiny model with a 16 × 16 patch size and a ResNet18 model with the final fully connected layer replaced by a binary classification head were used. Both models were initialized with ImageNet-pretrained weights and fine-tuned on the generated connectivity images. To handle class imbalance within each fold, weighted cross-entropy loss was applied. Let n_c denote the number of training samples in class c. The class weight was computed as:

$$w_c = \frac{1}{n_c} \tag{20}$$

The weighted cross-entropy loss is defined as:

$$\mathcal{L} = -\sum_{c=0}^{1} w_c \; y_c \log(\hat{p}_c) \tag{21}$$

where y_c is the ground-truth one-hot label and $\hat{p}_c$ is the predicted probability for class c.

TABLE II. TRAINING SETTINGS FOR THE DEEP LEARNING CLASSIFICATION SCHEMES

Setting	Single-band Schemes	Fusion Schemes
Models	DeiT-tiny, ResNet18	DeiT-tiny, ResNet18
Input size	224 × 224	224 × 224
Epochs	25	25
Batch size	128	128
Optimizer	RMSprop	RMSprop
Learning rate	1×10^{-6}	5×10^{-6}
Weight decay	1×10^{-6}	1×10^{-6}
Loss function	Weighted cross-entropy	Weighted cross-entropy
Pretrained weights	ImageNet	ImageNet
Cross-validation	8-fold stratified subject-wise cross-validation	8-fold stratified subject-wise cross-validation
Model selection	No best-weight selection; final epoch weights used	No best-weight selection; final epoch weights used
Augmentation	No geometric augmentation	No geometric augmentation

All models were trained for 25 epochs using RMSprop with a batch size of 128 and a weight decay of 1×10^{-6}. The learning rate was set to 1×10^{-6} for single-band schemes and 5×10^{-6} for fusion schemes. To reduce overfitting risk in the limited-subject setting, pretrained visual backbones, weight decay, weighted cross-entropy, preservation of the original PLV matrix topology, and fixed-epoch evaluation were used. No validation-based checkpoint selection or test-fold-based best-weight

selection was performed; instead, the final epoch weights were evaluated on the held-out test subjects. Subject-wise separation was maintained as described in Section II.F, ensuring that all window-derived images from the same subject remained within the same fold. The full training settings are summarized in Table II.

Overall, the deep learning stage comprised eight classification schemes: theta-DeiT, alpha-DeiT, beta-DeiT, theta-ResNet18, alpha-ResNet18, beta-ResNet18, fused-DeiT, and fused-ResNet18. This design enabled comparison across frequency bands, model families, and single-band versus multi-band fusion representations.

I. Performance Evaluation and Statistical Analysis

Classification performance was evaluated using standard confusion-matrix-based metrics, namely accuracy, precision, sensitivity, specificity, and F1-score, which are commonly used in binary medical and neurodevelopmental classification tasks [3, 11, 14, 21]. Let TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. The metrics were computed for each fold as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{22}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{23}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}, \tag{24}$$

$$\text{Specificity} = \frac{TN}{TN + FP}, \tag{25}$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \tag{26}$$

Accuracy measures overall classification correctness, precision reflects the reliability of ASD-positive predictions, sensitivity measures the ability to correctly identify ASD samples, specificity measures the correct identification of Normal samples, and F1-score provides a balanced summary of precision and sensitivity, particularly under class imbalance [11, 14]. All metrics were computed independently for each fold and summarized using the mean and standard deviation across the 8 folds.

Statistical analysis was performed using fold-wise paired accuracy values obtained from the subject-wise 8-fold cross-validation protocol. Because all classification schemes used the same subject-level fold partitioning, paired comparison was appropriate for controlling between-fold variability in this small-sample EEG classification setting [3, 36, 40]. Given the limited number of folds and the bounded nature of classification accuracy, a nonparametric Wilcoxon signed-rank test was used to compare the fusion schemes with their corresponding best-performing single-band baselines. Statistical significance was assessed at $\alpha = 0.05$ [3, 36].

The primary statistical comparison was conducted between each fusion scheme and its strongest single-band baseline. Since the beta band achieved the highest mean accuracy among the single-band schemes for both DeiT and ResNet18, fusion-DeiT was primarily compared with beta-DeiT, and fusion-ResNet18 was primarily compared with beta-ResNet18. This comparison provides the most conservative test of whether multi-band fusion adds discriminative value beyond the best individual EEG band.

For each fold f, the paired difference was defined as:

$$d_f^{\text{fusion}} = A_f^{\text{fusion}} - A_f^{\text{best-single}} \tag{27}$$

where A_f^{fusion} denotes the fold-wise accuracy of the fusion scheme and $A_f^{best\text{-}single}$ denotes the fold-wise accuracy of the corresponding best single-band scheme. Thus, the inferential analysis specifically tested whether the subject-wise fold performance of the fusion representation was significantly higher than that of the strongest single-band baseline.

To further address the contribution of the other EEG bands, additional post-hoc comparisons were conducted between each fusion scheme and the theta-, alpha-, and beta-band schemes from the same model family. Holm-adjusted p-values were reported to account for multiple comparisons [41]. Therefore, the statistical analysis included both the conservative primary comparison against the best single-band baseline and the extended post-hoc comparison against all evaluated EEG bands.

In addition to fold-wise window-derived image-level evaluation, a subject-level aggregation analysis was conducted as an additional robustness check for the best-performing fusion schemes. This analysis was performed only on the held-out test subjects in each fold. For each test subject, the predicted ASD probabilities from all window-derived connectivity images belonging to that subject were averaged to obtain a single subject-level probability. Let $p_s^{(m)}$ denote the predicted ASD probability for the m-th test window of subject s, and let M_s denote the number of test windows from that subject. The subject-level ASD probability was computed as:

$$P_s = \frac{1}{M_s} \sum_{m=1}^{M_s} p_s^{(m)} \tag{28}$$

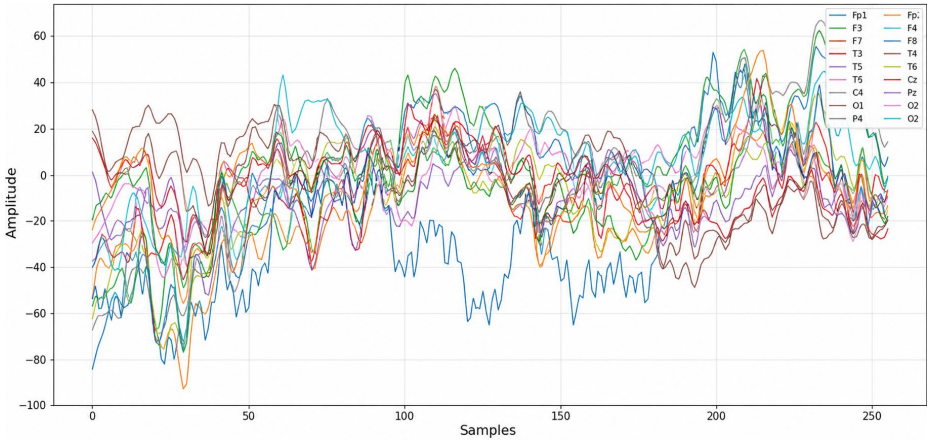
The final subject-level label was assigned as ASD when $P_s \geq 0.5$, and as Normal otherwise. The true-class probability was defined as the aggregated probability assigned to the correct class of each held-out subject.

III. Results and Discussion

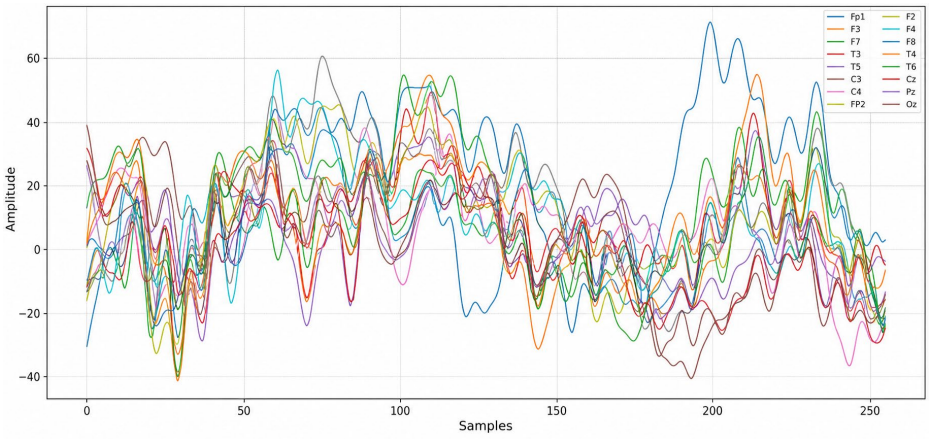
This section presents the experimental results of the proposed framework for ASD classification using multi-band EEG connectivity matrix images, including preprocessing results, connectivity image characteristics, classification performance, statistical comparisons, and discussion of the main findings.

A. Preprocessing Results

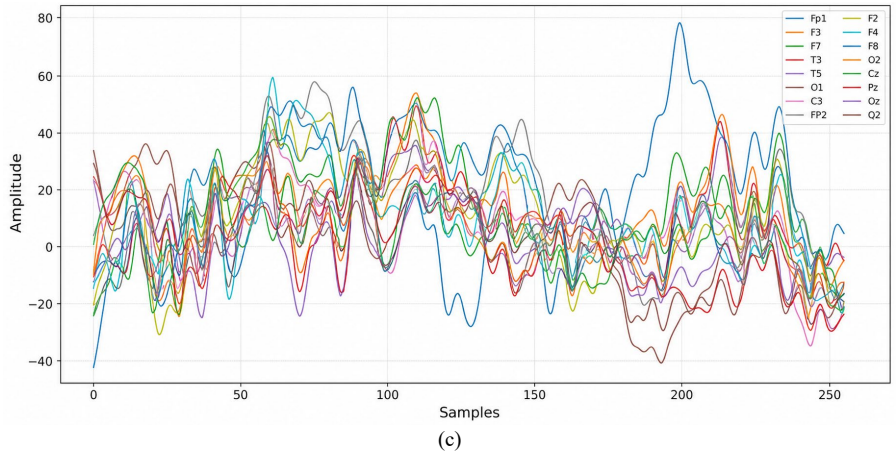
The preprocessing stage was designed to improve signal quality before frequency-band decomposition, PLV computation, and image-based classification. In the present study, preprocessing consisted of Butterworth bandpass filtering followed by ICA-based artifact attenuation. Fig. 4 presents representative examples of the EEG signals at three stages, namely the raw recordings, the signals after Butterworth filtering, and the signals after the combined Butterworth and ICA preprocessing pipeline.



(a)



(b)



(c)

Fig. 4. Representative 1-second multichannel EEG segment from an ASD subject at three preprocessing stages: (a) raw signal, showing baseline drift and irregular amplitude fluctuations; (b) after Butterworth bandpass filtering, showing reduced low-frequency drift and high-frequency noise; and (c) after combined Butterworth filtering and ICA, showing further attenuation of residual artifact-related components.

Fig. 4 illustrates the progressive improvement in signal quality prior to PLV connectivity estimation. Visual inspection of the raw EEG signals indicates that the original recordings contain substantial low-frequency fluctuation, abrupt amplitude variation, and inter-channel irregularity, which are typical characteristics of resting-state EEG contaminated by non-neural components such as baseline drift, movement-related activity, and physiological artifacts. Although the raw signals still preserve underlying neural oscillatory activity,

the waveform structure appears relatively unstable, making direct downstream analysis less reliable.

After Butterworth bandpass filtering, the EEG signals become visibly more regular and constrained within a narrower amplitude range. In particular, slow baseline drift is reduced and high-frequency fluctuations appear less prominent, indicating that the 0.5–45 Hz filtering stage successfully suppresses non-informative spectral components outside the target neural range. This visual improvement is important because subsequent band decomposition and phase-based connectivity analysis depend on signals with preserved oscillatory structure and reduced broadband contamination.

The application of ICA after Butterworth filtering further refines the multichannel EEG recordings. Compared with the filtered-only signals, the Butterworth + ICA signals exhibit smoother channel trajectories and reduced dominance of irregular transient fluctuations in specific channels. The overall inter-channel pattern becomes more coherent while preserving subject-specific waveform variability. This result is consistent with the intended role of ICA in attenuating residual artifact-related components that may not be fully removed by frequency filtering alone.

From a signal-processing perspective, the progression shown in Fig. 4 suggests that the combined preprocessing pipeline improves the suitability of the EEG recordings for the subsequent analytical stages. Butterworth filtering provides an initial spectral cleanup, whereas ICA contributes additional artifact suppression at the component level. As a result, the signals entering the frequency-band decomposition stage are visually cleaner and more structured than the raw recordings. This is particularly relevant for the proposed framework because PLV-based connectivity estimation is sensitive to phase behavior, and excessive artifact contamination may distort the synchronization patterns used to build the connectivity matrix images.

Fig. 4 should be interpreted as a qualitative illustration of signal-quality improvement rather than as stand-alone evidence of diagnostic group separation. Panel (a) shows the raw EEG segment with visible baseline drift and irregular amplitude variation, panel (b) shows the signal after Butterworth filtering with reduced slow drift and high-frequency fluctuation, and panel (c) shows the signal after Butterworth + ICA with further attenuation of residual artifact-related components. Overall, the progression across the three panels indicates a reduction of non-neural or artifact-related variability while preserving the oscillatory structure required for frequency decomposition and PLV-based connectivity estimation.

To complement the qualitative signal inspection in Fig. 4, a preprocessing ablation experiment was conducted to quantify the contribution of each preprocessing stage. The experiment was performed using the alpha-band DeiT setting as a representative classification configuration. Table III reports the alpha-band DeiT classification accuracy under three input conditions: raw-input condition without global preprocessing, global Butterworth-only condition, and the complete Butterworth + ICA preprocessing pipeline.

TABLE III. PREPROCESSING ABLATION RESULTS USING THE ALPHA-BAND DEIT SETTING

Preprocessing Condition	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Mean ± Std
Raw-input condition	0.8874	0.9081	0.9024	0.9325	0.9099	0.9099	0.9041	0.9361	0.9113 ± 0.0149
Global Butterworth only	0.9118	0.9362	0.9343	0.9400	0.9381	0.9343	0.9305	0.9492	0.9343 ± 0.0100
Butterworth + ICA	0.9579	0.9707	0.9780	0.9615	0.9579	0.9707	0.9780	0.9579	0.9666 ± 0.0083

The raw-input condition achieved a mean accuracy of 0.9113 ± 0.0149, indicating that the alpha-band PLV connectivity images still contained discriminative information even without the complete preprocessing pipeline. After applying the global 0.5–45 Hz Butterworth bandpass filter, the mean accuracy increased to 0.9343 ± 0.0100, corresponding to an improvement of 0.0230 over the raw-input condition. This improvement suggests that attenuating low-frequency drift and high-frequency noise produced a more stable signal representation for PLV-based connectivity estimation.

When ICA-based artifact attenuation was added after Butterworth filtering, the mean accuracy further increased to 0.9666 ± 0.0083. This represents an additional improvement of 0.0323 over the global Butterworth-only condition and 0.0553 over the raw-input condition. The reduction in standard deviation from 0.0149 to 0.0100 and then to 0.0083 also indicates more stable fold-wise performance. These results show that Butterworth filtering and ICA-based artifact attenuation contributed incrementally to preparing cleaner and more discriminative EEG representations for alpha-band PLV connectivity image classification.

B. Visual Characteristics of Connectivity Matrix Images

Fig. 5 illustrates the PLV-based connectivity matrix images generated from representative ASD and Normal samples across the theta, alpha, and beta frequency bands. Each image corresponds to a 16 × 16 PLV matrix derived from one EEG window. The bright diagonal elements represent self-connectivity fixed at the maximum PLV value, whereas the off-diagonal elements represent pairwise inter-channel phase synchronization. Therefore, variations in color intensity and spatial texture reflect differences in functional connectivity patterns among EEG channel pairs.

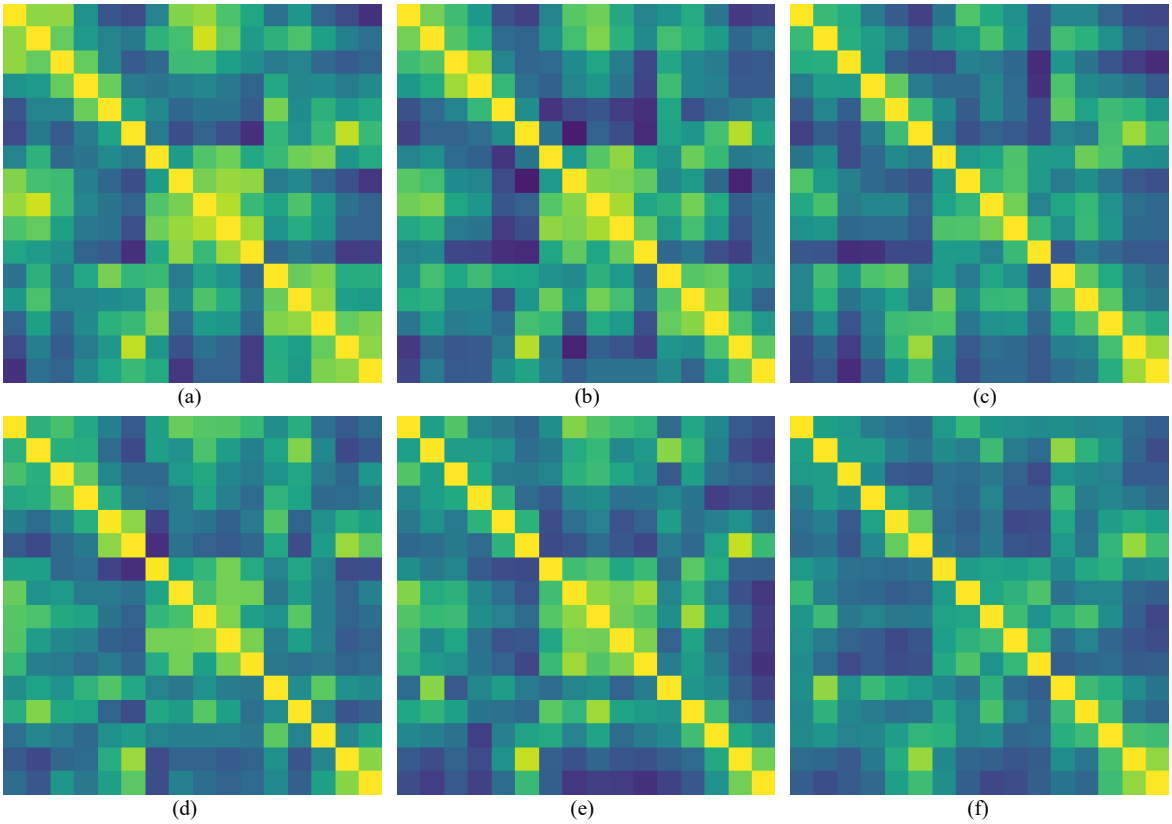


Fig. 5. Representative PLV-based 16 × 16 connectivity matrix images for ASD and Normal samples across theta, alpha, and beta bands. (a) ASD-Theta; (b) ASD-Alpha; (c) ASD-Beta; (d) Normal-Theta; (e) Normal-Alpha; (f) Normal-Beta.

Across frequency bands, the connectivity images show distinct visual characteristics. The theta-band images display broader intermediate-intensity regions, suggesting more distributed synchronization patterns. The alpha-band images exhibit more apparent local block-like variations, indicating localized changes in phase synchronization. The beta-band images show clearer localized intensity contrasts and more structured connectivity patterns compared with the theta and alpha representations. These band-dependent differences support the motivation for evaluating theta, alpha, and beta connectivity images separately and subsequently combining them through multi-band fusion.

A visual comparison between representative ASD and Normal samples suggests subtle differences in the spatial organization of PLV connectivity patterns. In particular, the beta-band representations appear more visually distinctive in terms of localized synchronization and intensity variation, which is consistent with the subsequent classification results showing that beta-band schemes achieved the strongest single-band performance. However, Fig. 5 should be interpreted as qualitative visual support rather than stand-alone statistical evidence of diagnostic separation. The primary quantitative evidence is provided by the classification results and statistical analysis, which further demonstrate the discriminative value of beta-band connectivity images and the additional benefit of theta-alpha-beta fusion.

C. Single-band Classification Results

The classification performance of the six single-band schemes is summarized in Table IV. These experiments were designed to evaluate the discriminative contribution of each frequency band when processed independently and classified using either the DeiT or ResNet18 architecture. Overall, the results show that all single-band schemes achieved strong classification performance, with mean accuracies above 0.93 across all settings. However, clear differences were observed both across frequency bands and across model architectures.

TABLE IV. CLASSIFICATION ACCURACY OF THE 6 SINGLE-BAND SCHEMES ACROSS THE 8 SUBJECT-WISE FOLDS

Scheme	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Mean ± Std
Alpha-DeiT	0.9579	0.9707	0.9780	0.9615	0.9579	0.9707	0.9780	0.9579	0.9666 ± 0.0083
Theta-DeiT	0.9670	0.9487	0.9615	0.9579	0.9487	0.9597	0.9670	0.9487	0.9574 ± 0.0074
Beta-DeiT	**0.9872**	**0.9835**	**0.9817**	**0.9780**	**0.9817**	**0.9817**	**0.9835**	**0.9762**	**0.9817 ± 0.0032**
Alpha-ResNet18	0.9432	0.9542	0.9524	0.9524	0.9359	0.9652	0.9725	0.9451	0.9526 ± 0.0111
Theta-ResNet18	0.9542	0.9377	0.9286	0.9359	0.9469	0.9286	0.9414	0.9304	0.9380 ± 0.0086
Beta-ResNet18	0.9560	0.9725	0.9652	0.9780	0.9670	0.9597	0.9799	0.9725	0.9689 ± 0.0079

Among the three frequency bands, the beta-band representations consistently produced the strongest results for both classifiers. In the transformer-based setting, the beta-band DeiT scheme achieved the highest mean

accuracy of 0.9817 ± 0.0032, outperforming both the alpha-band DeiT scheme, which reached 0.9666 ± 0.0083, and the theta-band DeiT scheme, which yielded 0.9574 ± 0.0074. This indicates that the beta-band connectivity images contained more discriminative information for ASD classification than the lower-frequency bands when evaluated using the DeiT architecture.

A similar pattern was observed for ResNet18. The beta-band ResNet18 scheme achieved the highest mean accuracy of 0.9689 ± 0.0079, followed by the alpha-band scheme with 0.9526 ± 0.0111 and the theta-band scheme with 0.9380 ± 0.0086. These results suggest that beta-band synchronization patterns provide the most informative single-band representation among the three evaluated frequency ranges.

When the two architectures are compared within the same band, DeiT consistently outperformed ResNet18. In the alpha band, DeiT improved the mean accuracy from 0.9526 to 0.9666. In the theta band, the improvement was from 0.9380 to 0.9574. In the beta band, the advantage of DeiT remained evident, increasing the mean accuracy from 0.9689 to 0.9817. These results suggest that the transformer-based model was better able to capture the spatial dependencies embedded in the connectivity matrix images than the convolution-based baseline.

In addition to the mean accuracy, the standard deviation across folds provides insight into performance stability. The beta-band DeiT scheme exhibited not only the highest mean accuracy but also the lowest standard deviation, indicating both strong discriminative capability and consistent performance across folds. In contrast, the alpha-band ResNet18 scheme showed the largest standard deviation among the six schemes, suggesting greater sensitivity to fold-specific variation. Thus, the beta-band DeiT configuration emerged as the most reliable single-band setting in the present study.

Taken together, the single-band results reveal two main findings. First, among the three frequency bands, beta-band connectivity images provided the strongest individual representation for ASD classification. Second, DeiT consistently achieved better performance than ResNet18 across all bands, suggesting that transformer-based visual modeling may be more effective for PLV-derived EEG connectivity images. These observations motivate the subsequent evaluation of multi-band fusion, where complementary information from theta, alpha, and beta representations is integrated within a unified classification framework.

D. Multi-band Fusion Classification Results

The classification performance of the two multi-band fusion schemes is summarized in Table V. In this setting, the theta, alpha, and beta connectivity images corresponding to the same subject and window index were combined into a unified three-channel representation before being classified using either DeiT or ResNet18. This experiment was designed to evaluate whether integrating complementary information from multiple frequency bands could improve ASD classification performance beyond that achieved by the best single-band schemes.

TABLE V. MULTI-BAND FUSION CLASSIFICATION ACCURACY ACROSS THE 8 SUBJECT-WISE FOLDS

Scheme	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Mean ± Std
Fusion-DeiT	0.9927	0.9945	0.9835	0.9963	0.9908	0.9908	0.9872	0.9817	0.9897 ± 0.0048
Fusion-ResNet18	0.9908	0.9927	0.9822	0.9872	0.9835	0.9835	0.9863	0.9817	0.9860 ± 0.0041

The results show that both fusion schemes achieved higher mean accuracy than any of the individual single-band configurations. The fusion-based DeiT model obtained the best overall performance, with a mean accuracy of 0.9897 ± 0.0048, whereas the fusion-based ResNet18 model achieved a mean accuracy of 0.9860 ± 0.0041. These results indicate that combining theta-, alpha-, and beta-band connectivity information yielded a clear performance gain over the separate band-specific representations.

When compared with the strongest single-band results reported in Table IV, the advantage of fusion becomes evident. For DeiT, the fusion scheme improved the mean accuracy from 0.9817 ± 0.0032 in the beta-band setting to 0.9897 ± 0.0048. Likewise, for ResNet18, the fusion strategy increased the mean accuracy from 0.9689 ± 0.0079 in the beta-band setting to 0.9860 ± 0.0041. These improvements suggest that the fusion representation was able to exploit complementary synchronization patterns distributed across multiple frequency ranges, thereby enhancing the separability between ASD and Normal samples.

A comparison between the two fusion architectures further shows that DeiT retained a modest performance advantage over ResNet18. Although both models performed exceptionally well under the fusion setting, DeiT produced a slightly higher mean accuracy. This finding is consistent with the single-band results, where DeiT also outperformed ResNet18 across all three bands. The result suggests that the transformer-based architecture remained more effective in modeling the spatial dependencies embedded in the fused connectivity images.

In terms of stability, both fusion schemes exhibited low standard deviation across folds, indicating consistent performance under subject-wise cross-validation. Notably, the fusion-based ResNet18 model showed a slightly lower standard deviation than fusion-based DeiT, whereas fusion-based DeiT still achieved the highest mean accuracy overall. This suggests that both architectures benefited from multi-band integration, although DeiT offered the strongest overall discriminative performance.

Taken together, the fusion results reveal that integrating theta, alpha, and beta connectivity matrix images is more effective than relying on any single-band representation alone. The superior performance of the fusion schemes

indicates that the three bands contain complementary information relevant to ASD classification, and that combining them within a unified image-based representation can strengthen model performance. These findings support the central premise of the proposed framework and motivate the statistical comparisons presented in the following subsection.

E. Statistical Analysis Results

To evaluate whether multi-band fusion significantly improved performance over single-band representations, Wilcoxon signed-rank tests were applied to the fold-wise paired accuracy values. The primary comparison was made against the beta band because beta produced the highest single-band accuracy for both DeiT and ResNet18. Therefore, this comparison represents the strictest single-band baseline rather than a selective comparison.

In addition to the primary comparison, post-hoc comparisons were conducted between each fusion scheme and all single-band schemes from the same model family, namely theta, alpha, and beta. Holm-adjusted *p*-values were reported to account for multiple comparisons. The corresponding statistical outcomes are reported in Table VI.

TABLE VI. STATISTICAL COMPARISON BETWEEN FUSION AND SINGLE-BAND SCHEMES

Comparison	Mean Diff.	*p*-value	Holm-adjusted *p*-value	RBC	Result
Fusion-DeiT vs Theta-DeiT	+0.0323	0.0078	0.0469	1.000	Sig.
Fusion-DeiT vs Alpha-DeiT	+0.0231	0.0078	0.0469	1.000	Sig.
Fusion-DeiT vs Beta-DeiT	+0.0080	0.0078	0.0469	1.000	Sig.
Fusion-ResNet18 vs Theta-ResNet18	+0.0480	0.0078	0.0469	1.000	Sig.
Fusion-ResNet18 vs Alpha-ResNet18	+0.0334	0.0078	0.0469	1.000	Sig.
Fusion-ResNet18 vs Beta-ResNet18	+0.0171	0.0078	0.0469	1.000	Sig.

The results showed that the fusion schemes significantly outperformed all single-band schemes for both model architectures. For DeiT, fusion-DeiT improved the mean accuracy over theta-DeiT, alpha-DeiT, and beta-DeiT by +0.0323, +0.0231, and +0.0080, respectively. For ResNet18, fusion-ResNet18 improved the mean accuracy over theta-ResNet18, alpha-ResNet18, and beta-ResNet18 by +0.0480, +0.0334, and +0.0171, respectively.

Although fusion was compared with all EEG bands in the post-hoc analysis, the comparison with beta remains the primary and most conservative comparison because beta was the strongest individual band for both architectures. These findings indicate that theta, alpha, and beta provide complementary information, while beta remains the strongest single-band representation. Thus, the statistical analysis supports the main claim that multi-band fusion provides additional discriminative value beyond any single EEG band.

F. Subject-level Robustness Analysis

Because the best-performing schemes achieved very high window-level classification accuracies, an additional subject-level aggregation analysis was conducted to examine whether the results remained consistent at the biological subject level. Using the aggregation procedure described in Section II.I, the predicted probabilities from all window-derived connectivity images belonging to each held-out test subject were averaged to obtain a single subject-level prediction. This analysis was performed only on the held-out test subjects in each fold.

As shown in Table VII, both fusion schemes correctly classified all 16 held-out subjects after subject-level aggregation. Fusion-DeiT achieved a subject-level accuracy of 100.00%, with a mean true-class probability of 98.42% and a minimum true-class probability of 94.31%. Fusion-ResNet18 also achieved a subject-level accuracy of 100.00%, with a mean true-class probability of 97.86% and a minimum true-class probability of 92.74%. These findings indicate that, within the evaluated dataset, the high window-level accuracies were consistent with subject-level predictions and were not driven only by isolated window-level decisions. However, the 100.00% subject-level accuracy should be interpreted cautiously. The difference between window-level and subject-level accuracy is due to the aggregation procedure: some individual windows may still be misclassified, but their effects can be outweighed when the predicted probabilities from all windows of the same held-out subject are averaged. Therefore, the subject-level result indicates that the dominant prediction tendency for each held-out subject was correct, not that every window-derived image was classified perfectly. Because only 16 subjects were available, this result should be regarded as within-dataset robustness evidence rather than proof of broad generalizability.

TABLE VII. SUBJECT-LEVEL AGGREGATION ANALYSIS OF THE FUSION SCHEMES

Scheme	Subject-level Accuracy (%)	Correct Subjects	Mean True-class Probability (%)	Min. True-class Probability (%)	Aggregation Rule
Fusion-DeiT	100.00	16/16	98.42	94.31	Average window probability
Fusion-ResNet18	100.00	16/16	97.86	92.74	Average window probability

G. Discussion of Findings

The findings of this study highlight three main points. First, among the single-band schemes, the beta-band connectivity images consistently yielded the strongest classification performance. Second, the multi-band fusion strategy significantly improved performance relative to the best single-band configuration. Third, the transformer-based DeiT model consistently outperformed the CNN-based ResNet18 model across both single-band and fusion settings. Taken together, these results indicate

that ASD-related EEG connectivity patterns can be effectively represented as PLV-based images and that both spectral selection and model architecture play important roles in classification performance.

The superior performance of the beta-band schemes suggests that beta-range phase synchronization contains highly discriminative information for distinguishing ASD from Normal subjects in the present dataset. Biologically, beta oscillations are commonly associated with sensorimotor regulation, maintenance of ongoing cortical states, top-down control, and long-range functional integration. Therefore, stronger discriminability in beta-band PLV connectivity may indicate that ASD-related differences in this dataset are more clearly reflected in the consistency of inter-channel phase synchronization within networks supporting sensorimotor coordination and distributed cortical communication. This interpretation is consistent with recent resting-state EEG findings showing atypical beta-range activity in ASD and with broader evidence that ASD is associated with altered functional connectivity and neural synchronization [1, 3, 42, 43]. From a neurodevelopmental perspective, abnormal beta-band synchronization may reflect atypical excitation-inhibition balance, cortical hyperexcitability, and altered integration between local and distributed brain networks, which have been frequently discussed as candidate mechanisms underlying sensory processing, behavioral regulation, and social-communication difficulties in ASD [2, 42, 43]. In this study, PLV captures the stability of phase relationships between EEG channel pairs; thus, the superior beta-band performance suggests that beta-range inter-channel synchronization patterns may provide a biologically meaningful representation of ASD-related network organization. Nevertheless, this interpretation remains hypothesis-generating because the present classification framework does not directly establish causal neurophysiological mechanisms.

The results also demonstrate the importance of integrating information across multiple frequency bands. While the beta band provided the strongest single-band representation, the fusion schemes achieved even higher performance for both architectures, and the Wilcoxon signed-rank tests confirmed that these improvements were statistically significant. This suggests that theta-, alpha-, and beta-band connectivity images contain complementary rather than redundant information. From a representation-learning perspective, the fusion strategy enables the model to access multiple views of functional brain organization within a single input, thereby strengthening class separability. This finding is important because ASD-related neural abnormalities are unlikely to be confined to a single oscillatory range. Instead, the disorder may involve multi-scale disruptions in functional coordination, making band integration a more informative strategy than isolated band analysis.

Beyond the reported classification performance, the main methodological implication of the proposed framework is that EEG functional connectivity can be reformulated as a structured image-fusion problem. Instead of treating frequency bands as independent feature sets only, the proposed approach preserves band-specific PLV connectivity patterns and combines them at the image-channel level using subject- and window-aligned fusion. This design allows standard visual deep learning models to jointly learn complementary connectivity patterns across multiple oscillatory rhythms. Therefore, the study contributes a reproducible representation-level fusion strategy for EEG-based ASD classification, while the use of DeiT and ResNet18 serves to evaluate how different visual learning paradigms respond to the proposed representation.

Another notable finding is the consistent advantage of DeiT over ResNet18. Across all single-band schemes and the fusion setting, DeiT achieved higher mean accuracy than ResNet18. In the single-band experiments, DeiT improved the mean accuracy from 0.9526 to 0.9666 in the alpha band, from 0.9380 to 0.9574 in the theta band, and from 0.9689 to 0.9817 in the beta band. In the fusion setting, DeiT also achieved the highest overall mean accuracy of 0.9897 ± 0.0048, compared with 0.9860 ± 0.0041 for ResNet18.

These results suggest that the strength of DeiT lies in its ability to capture global and long-range dependencies within PLV-derived connectivity matrix images. This is particularly relevant because each pixel or region in a connectivity image represents relationships between EEG channel pairs, and diagnostically useful information may depend on distributed inter-channel synchronization patterns rather than only local image textures. ResNet18 also achieved strong and stable performance, especially in the fusion setting, indicating that convolutional residual learning remains effective for extracting local and block-like connectivity patterns. However, the consistently higher performance of DeiT suggests that attention-based modeling provides an additional advantage when EEG functional connectivity is represented as an image.

Although a separate model-attribution experiment, such as Grad-CAM (Class Activation Mapping) or attention rollout, was not performed in the present study, the proposed PLV connectivity image representation provides representation-level interpretability. Each matrix location corresponds to a fixed EEG channel-pair relationship, and its intensity represents the strength of PLV-based phase synchronization between the corresponding channels. Therefore, the model input is not an arbitrary image, but a neurophysiologically meaningful representation of inter-channel synchronization. Accordingly, the stronger beta-band and fusion performance suggest that discriminative information was more clearly represented in beta-range and multi-band synchronization patterns. However, this interpretation remains qualitative; future work should incorporate model-level XAI methods, such as Grad-CAM for CNNs and attention rollout for transformers, to identify the specific channel-pair regions that contribute most strongly to ASD classification.

To further contextualize the proposed framework, Table VIII compares the present study with previous EEG-based ASD classification studies that reported classification

accuracy. The comparison includes studies using Short-Time Fourier Transform (STFT) spectrogram EEG images, augmented EEG-as-image windows, scalogram representations, and the proposed multi-band PLV connectivity matrix images. Because the compared studies used different datasets, subject compositions, preprocessing procedures, feature representations, and validation protocols, the comparison should be interpreted as a methodological and contextual comparison rather than a direct one-to-one performance ranking.

TABLE VIII. COMPARISON WITH PREVIOUS EEG-BASED ASD CLASSIFICATION STUDIES

Study	Dataset/Subjects	Method	Feature/Input	Validation Protocol	Accuracy (%)
Tawhid *et al.* [44]	KAU EEG dataset, 16 subjects	CNN-based DL model	STFT spectrogram EEG images	70% training, 15% validation, 15% testing	99.15
Ardakani *et al.* [9]	EEG ASD dataset	2D-DCNN with data augmentation	Augmented EEG-as-image windows	Window-level classification and joined-window subject decision	88.29/100.00
Torghabeh *et al.* [4]	34 ASD and 11 Normal EEG samples	Hybrid CNN-LSTM deep transfer learning	Scalogram EEG images	Older-than-5 subjects for training; younger subjects for validation	99.50/98.43
Present study	KAU EEG dataset, 16 subjects	DeiT/ResNet18	Multi-band PLV connectivity matrix images	Subject-wise 8-fold cross-validation	98.97/98.60

The studies summarized in Table VIII indicate that recent EEG-based ASD classification research has commonly used image-like EEG representations, including STFT spectrograms, augmented EEG-as-image windows, and scalogram images. These prior works support the usefulness of transforming EEG signals into visual representations for deep learning-based ASD classification.

In contrast to these approaches, the present study focuses on phase-based functional connectivity by converting theta-, alpha-, and beta-band PLV matrices into connectivity images and integrating them through subject- and window-aligned multi-band fusion. Therefore, the comparison suggests that the present framework complements previous EEG image-based ASD studies by emphasizing multi-band connectivity representation rather than only signal morphology or time-frequency energy patterns. The main distinction of the proposed framework lies in its representation-level fusion strategy and its evaluation under subject-wise 8-fold cross-validation with subject-level robustness analysis.

In summary, the proposed framework provides a unified pipeline that connects EEG preprocessing, frequency decomposition, PLV-based connectivity modeling, image-level multi-band fusion, and deep visual classification. This integration shows how multichannel EEG synchronization patterns can be translated into structured image representations suitable for modern visual learning models. The results further support the use of subject- and window-aligned spectral fusion as a representation strategy for EEG-based ASD classification, although its generalizability should be confirmed in larger independent datasets.

At the same time, several limitations should be acknowledged. The small number of biological subjects remains an important limitation. The 16-subject dataset is sufficient only for evaluating the proposed framework as an internal proof-of-concept under strict subject-wise cross-validation, not for establishing broad clinical generalizability. Although windowing generated multiple EEG segments for model training, all windows from the same subject were kept within the same fold to prevent subject-level data leakage. The use of subject-wise splitting, transfer learning, weight decay, weighted cross-entropy, fixed-epoch evaluation, and subject-level aggregation was intended to reduce leakage and overfitting risk. The high fusion and subject-level aggregation performance may partly reflect the controlled and limited structure of the evaluated dataset; therefore, overoptimistic performance cannot be fully ruled out without external validation. Although ImageNet pretraining helped stabilize learning in this limited-data setting, future work should compare it with random initialization and EEG-specific self-supervised pretraining for PLV connectivity images. Accordingly, validation on larger, independent, gender-balanced, and preferably multi-center datasets is required before the proposed approach can be considered clinically robust. Another dataset-related limitation is the absence of female participants in the control group. Because sex-related effects cannot be fully separated from diagnosis-related effects in the present dataset, future validation should use sex-balanced ASD and control cohorts. In addition, the source dataset documentation does not provide detailed information on the diagnostic instruments used for ASD confirmation, which restricts the clinical interpretability of the resulting classification performance. Another limitation is that the present work focused on PLV as the connectivity measure; other functional connectivity metrics may reveal complementary relationships that were not explored here.

Despite these limitations, the present results provide encouraging evidence that multi-band EEG connectivity image fusion is a promising direction for further investigation in EEG-based ASD classification. The combination of subject-wise evaluation, strong single-band performance, statistically validated fusion improvement, and consistent gains from transformer-based modeling indicates that the proposed framework captures discriminative EEG connectivity characteristics within the evaluated dataset. Future work may extend this approach by evaluating additional connectivity measures, incorporating larger multi-center datasets, and exploring more advanced fusion strategies to further improve robustness and interpretability.

IV. Conclusion

This study proposed a framework for ASD classification using multi-band EEG connectivity matrix images derived from resting-state signals. The framework integrated Butterworth bandpass filtering, ICA-based artifact attenuation, frequency-band decomposition, PLV-based functional connectivity analysis, connectivity image generation, and deep learning classification using DeiT and ResNet18. The results showed that all single-band schemes achieved high classification performance, with the beta band providing the strongest individual representation for both architectures. Furthermore, the multi-band fusion strategy produced the best overall performance, and statistical analysis confirmed that fusion significantly outperformed the best single-band baseline for both models.

Overall, within the evaluated dataset, these findings demonstrate that multi-band EEG connectivity image fusion is a promising strategy for ASD classification and that transformer-based modeling offers an advantage over the CNN-based baseline for PLV-derived connectivity representations. The proposed framework therefore provides a meaningful integration of EEG signal processing, functional connectivity analysis, and image-based deep learning within a unified classification pipeline. Future work may extend this approach by evaluating additional connectivity measures, validating the framework on larger and more diverse datasets, and exploring more advanced fusion and transformer-based architectures to improve robustness and interpretability.

Data Availability

The EEG dataset used in this study was obtained from King Abdulaziz University. The data are available from the data owner upon reasonable request and with appropriate permission.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Z.Z. contributed to the conceptualization and supervision of the study. J.J. contributed to the methodology and study design. K.M.S. contributed to validation and interpretation of the results. A.R. contributed to data processing, experimental evaluation, and manuscript drafting. M.D. contributed to literature review and manuscript editing. All authors reviewed and approved the final version of the manuscript.

Acknowledgment

The authors would like to express their sincere appreciation to Universitas Syiah Kuala (USK) for the institutional support provided during the completion of this research.

References

[1] X. Tong *et al.*, "Symptom dimensions of resting-state electroencephalographic functional connectivity in autism," *Nature. Mental Health*, vol. 2, no. 3, pp. 287–298, 2024. doi: 10.1038/s44220-023-00195-w

[2] J. Shan *et al.*, "A scoping review of physiological biomarkers in autism," *Front. Neurosci.*, vol. 17, 1269880, 2023. doi: 10.3389/fnins.2023.1269880

[3] J. Li *et al.*, "Identification of autism spectrum disorder based on electroencephalography: A systematic review," *Comput. Biol. Med.*, vol. 170, 108075, 2024. doi: 10.1016/j.compbiomed.2024.108075

[4] F. Abedinzadeh Torghabeh, Y. Modaresnia, and M. H. Moattar, "Hybrid deep transfer learning-based early diagnosis of autism spectrum disorder using scalogram representation of electroencephalography signals," *Medical & Biological Engineering & Computing*, vol. 62, no. 2, pp. 495–503, 2023. doi: 10.1007/s11517-023-02959-z

[5] W. Prany *et al.*, "EEG resting-state functional connectivity: Evidence for an imbalance of external/internal information integration in autism," *Journal of Neurodevelopmental Disorders*, vol. 14, no. 1, 47, 2022. doi: 10.1186/s11689-022-09456-8

[6] L. Wan *et al.*, "Multimodal investigation of dynamic brain network alterations in autism spectrum disorder: Linking connectivity dynamics to symptoms and developmental trajectories," *Neuroimage*, vol. 302, 120895, 2024. doi: 10.1016/j.neuroimage.2024.120895

[7] S. N. S. Fakhari, F. Ghaderi, M. Tehrani-Doost, and N. M. Charkari, "EEG-based brain connectivity analysis in autism spectrum disorder: Unraveling the effects of bumetanide treatment," *Biomed. Signal Process. Control*, vol. 86, 105054, 2023. doi: 10.1016/j.bspc.2023.105054

[8] M. S. Kabir, S. Kurkin, G. Portnova, O. Martynova, Z. Wang, and A. Hramov, "Contrastive machine learning reveals in EEG resting-state network salient features specific to autism spectrum disorder," *Chaos Solitons Fractals*, vol. 185, 115123, 2024. doi: 10.1016/j.chaos.2024.115123

[9] H. A. Ardakani, M. Taghizadeh, and F. Shayegh, "Diagnosis of Autism Disorder Based on Deep Network Trained by Augmented EEG Signals.," *Int. J. Neural Syst.*, vol. 32, no. 11, pp. 2250046–2250046, 2022. doi: 10.1142/S0129065722500460

[10] J. Rogala *et al.*, "Enhancing autism spectrum disorder classification in children through the integration of traditional statistics and classical machine learning techniques in EEG analysis," *Sci. Rep.*, vol. 13, no. 1, 21748, 2023. doi: 10.1038/s41598-023-49048-7

[11] S. Y. Ke *et al.*, "Classification of autism spectrum disorder using electroencephalography in Chinese children: A cross-sectional retrospective study," *Front. Neurosci.*, vol. 18, 1330556, 2024. doi: 10.3389/fnins.2024.1330556

[12] Q. M. U. Din and A. K. Jayanthy, "Automated classification of autism spectrum disorder using EEG signals and convolutional neural networks," *Biomed. Eng. (Singapore).*, vol. 34, no. 2, 2250020, 2022. doi: 10.4015/S101623722250020X

[13] T. Tang *et al.*, "A hybrid graph network model for ASD diagnosis based on resting-state EEG signals," *Brain Res. Bull.*, vol. 206, no. 1, 110826, Jan. 2024. doi: 10.1016/j.brainresbull.2023.110826

[14] 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," *Comput. Methods Programs Biomed.*, vol. 250, 108196, 2024. doi: 10.1016/j.cmpb.2024.108196

[15] M. J. Alhaddad *et al.*, "Diagnosis Autism by Fisher Linear Discriminant Analysis FLDA via EEG," *International Journal of Bio-Science and Bio-Technology*, vol. 4, no. 2, pp. 45–53, 2012.

[16] M. Murias, S. J. Webb, J. Greenson, and G. Dawson, "Resting State Cortical Connectivity Reflected in EEG Coherence in Individuals with Autism," *Biol. Psychiatry*, vol. 62, no. 3, pp. 270–273, 2007. doi: 10.1016/j.biopsych.2006.11.012

[17] E. V. Orekhova *et al.*, "Excess of High Frequency Electroencephalogram Oscillations in Boys with Autism," *Biol. Psychiatry*, vol. 62, no. 9, pp. 1022–1029, 2007. doi: 10.1016/j.biopsych.2006.12.029

[18] R. Coben, A. R. Clarke, W. Hudspeth, and R. J. Barry, "EEG power and coherence in autistic spectrum disorder," *Clinical*

Neurophysiology, vol. 119, no. 5, pp. 1002–1009, 2008. doi: 10.1016/j.clinph.2008.01.013
[19] F. Fahmi, M. Melinda, P. D. Purnamasari, E. Elizar, and A. Rafiki, "Recognition of EEG Features in Autism Disorder Using SWT and Fisher Linear Discriminant Analysis," *Diagnostics*, vol. 15, no. 18, 2291, 2025. doi: 10.3390/diagnostics15182291
[20] M. Melinda, F. H. Juwono, I. K. A. Enriko, M. Oktiana, S. Mulyani, and K. Saddami, "Application of Continuous Wavelet Transform and Support Vector Machine for Autism Spectrum Disorder Electroencephalography Signal Classification," *Radioelectronic and Computer Systems*, vol. 3, no. 107, pp. 73–90, 2023. doi: 10.32620/reks.2023.3.07
[21] A. Chaddad, Y. Wu, R. Kateb, and A. Bouridane, "Electroencephalography Signal Processing: A Comprehensive Review and Analysis of Methods and Techniques," *Sensors*, vol. 23, no. 14, 6434, 2023. doi: 10.3390/S23146434
[22] S. Das *et al.*, "Machine learning approaches for electroencephalography and magnetoencephalography analyses in autism spectrum disorder: A systematic review," *Prog. Neuropsychopharmacol. Biol. Psychiatry*, vol. 123, 110705, 2023. doi: 10.1016/j.pnpbp.2022.110705
[23] S. Coelli *et al.*, "Selecting methods for a modular EEG pre-processing pipeline: An objective comparison," *Biomed. Signal Process. Control*, vol. 90, 105830, 2024. doi: 10.1016/J.BSPC.2023.105830
[24] A. Li, J. Feitelberg, A. P. Saini, R. Höchenberger, and M. Scheltienne, "MNE-ICALabel: Automatically annotating ICA components with ICLabel in Python," *J. Open Source Softw.*, vol. 7, no. 76, 4484, 2022. doi: 10.21105/joss.04484
[25] A. M. Judith, S. B. Priya, and R. K. Mahendran, "Artifact Removal from EEG signals using Regenerative Multi-Dimensional Singular Value Decomposition and Independent Component Analysis," *Biomed. Signal Process. Control*, vol. 74, 103452, 2022. doi: 10.1016/j.bspc.2021.103452
[26] M. Klug, T. Berg, and K. Gramann, "Optimizing EEG ICA decomposition with data cleaning in stationary and mobile experiments," *Sci. Rep.*, vol. 14, no. 1, 14119, 2024. doi: 10.1038/S41598-024-64919-3
[27] B. Kalita, N. Deb, and D. Das, "AnEEG: Leveraging deep learning for effective artifact removal in EEG data," *Scientific Reports*, vol. 14, no. 1, 24234, 2024. doi: 10.1038/s41598-024-75091-z
[28] F. Artoni and C. M. Michel, "How does Independent Component Analysis Preprocessing Affect EEG Microstates?" *Brain Topography*, vol. 38, no. 2, 26, 2025. doi: 10.1007/s10548-024-01098-4
[29] C. H. Chuang, K. Y. Chang, C. S. Huang, and T. P. Jung, "IC-U-Net: A U-Net-based Denoising Autoencoder Using Mixtures of Independent Components for Automatic EEG Artifact Removal," *Neuroimage*, vol. 263, 119586, 2022. doi: 10.1016/j.neuroimage.2022.119586
[30] 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, 389, 2023. doi: 10.1038/s41398-023-02681-2
[31] S. J. K. Barnes, M. Thomas, P. V. E. McClintock, and A. Stefanovska, "Theta and alpha connectivity in children with autism spectrum disorder," *Brain Commun.*, vol. 7, no. 2, fcaf084, 2025. doi: 10.1093/braincomms/fcaf084
[32] W. Xiong, L. Ma, and H. Li, "A general dual-pathway network for EEG denoising," *Front. Neurosci.*, vol. 17, 1258024, 2023. doi: 10.3389/fnins.2023.1258024
[33] Z. M. Wang, S. Li, J. Zhang, and C. Liang, "Emotion recognition based on phase-locking value brain functional network and topological data analysis," *Neural Comput. Appl.*, vol. 36, no. 14, pp. 7903–7922, 2024. doi: 10.1007/s00521-024-09479-3
[34] W. Liu, K. Jia, and Z. Wang, "Graph-based EEG approach for depression prediction: Integrating time-frequency complexity and spatial topology," *Front. Neurosci.*, vol. 18, 2024. doi: 10.3389/fnins.2024.1367212
[35] L. Cao *et al.*, "A Novel Deep Learning Method Based on an Overlapping Time Window Strategy for Brain-Computer Interface-Based Stroke Rehabilitation," *Brain Sci.*, vol. 12, no. 11, 1502, 2022. doi: 10.3390/brainsci12111502
[36] M. Elbattah, O. Ali Sadek Ibrahim, and G. Dequen, "Editorial: Improving autism spectrum disorder diagnosis using machine learning techniques," *Front. Neuroinform.*, vol. 18, 1529839, 2024. doi: 10.3389/fninf.2024.1529839
[37] G. Chiarion, L. Sparacino, Y. Antonacci, L. Faes, and L. Mesin, "Connectivity Analysis in EEG Data: A Tutorial Review of the State of the Art and Emerging Trends," *Bioengineering*, vol. 10, no. 3, 372, 2023. doi: 10.3390/BIOENGINEERING10030372
[38] S. Raveendran *et al.*, "Functional connectivity in EEG: A multiclass classification approach for disorders of consciousness," *Front. Neurosci.*, vol. 19, 1550581, 2025. doi: 10.3389/fnins.2025.1550581
[39] Z. J. Peya, M. A. Maria, S. I. Hossain, M. A. H. Akhand, and N. Siddique, "Autism Spectrum Disorder Detection Using Prominent Connectivity Features from Electroencephalography," *Int. J. Neural Syst.*, vol. 35, no. 3, 2550011, 2025. doi: 10.1142/S012906572550011X
[40] I. Nafisah *et al.*, "Deep learning-based feature selection for detection of autism spectrum disorder," *Front. Artif. Intell.*, vol. 8, 1594372, 2025. doi: 10.3389/frai.2025.1594372
[41] S. Holm, "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, 1979. http://www.jstor.org/stable/4615733
[42] M. Proteau-Lemieux *et al.*, "Specific EEG resting state biomarkers in FXS and ASD," *J. Neurodev. Disord.*, vol. 16, no. 1, 53, 2024. doi: 10.1186/s11689-024-09570-9
[43] X. Geng *et al.*, "Abnormalities of EEG Functional Connectivity and Effective Connectivity in Children with Autism Spectrum Disorder," *Brain Sciences*, vol. 13, no. 1, 130, 2023. doi: 10.3390/BRAINSCI13010130
[44] 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, 2021. doi: 10.1371/journal.pone.0253094

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).