Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 3, July 2026, pp: 1239-1259 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

Temporal Deep Learning with Multiscale Principal Component Features for Autism Classification from Electroencephalographic Signals

Muliyadi Muliyadi[1], Melinda Melinda[2], Yuwaldi Away[2], Syahrul Gazali[3,4], Aufa Rafiki[2], and W. K. Wong[5]

[1] Doctoral Program in Engineering Science, Postgraduate School, Universitas Syiah Kuala, Banda Aceh, Indonesia.
[2] Department of Electrical and Computer Engineering, Faculty of Engineering, Universitas Syiah Kuala, Banda Aceh, Indonesia.
[3] Department of Neurology, Faculty of Medicine, Universitas Syiah Kuala, Banda Aceh, Indonesia.
[4] Department of Neurology, RSUD dr. Zainoel Abidin, Banda Aceh, Indonesia.
[5] Department of Electrical and Computer Engineering, Faculty of Engineering and Sciences, Curtin University Malaysia, Miri, Malaysia.

Corresponding author: Melinda Melinda (e-mail: melinda@usk.ac.id), **Author(s) Email**: Muliyadi Muliyadi (e-mail: muliyadi@pnl.ac.id), Yuwaldi Away (e-mail: yuwaldi@usk.ac.id), Syahrul Gazali (e-mail: syahrulsps@usk.ac.id), Aufa Rafiki (e-mail: aufa35@mhs.usk.ac.id), W. K. Wong (e-mail: weikitt.w@ieee.org)

Abstract **Autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental condition whose early identification remains challenging because clinical assessment still relies heavily on behavioral observation and expert judgment. Electroencephalography (EEG) offers a noninvasive approach for capturing neural dynamics. Still, EEG-based ASD classification remains difficult because of signal nonstationarity, limited sample sizes, and the risk of subject-identity leakage. This study proposes a comparative temporal deep learning framework for EEG-based ASD classification by evaluating principal component analysis (PCA) and multiscale principal component analysis (MS-PCA) as feature representations combined with a recurrent neural network with bidirectional long short-term memory (RNN-BiLSTM) and a temporal convolutional network with self-attention (TCN-SA). Resting-state eyes-open EEG signals were acquired from only 10 participants, consisting of 5 individuals with ASD and 5 typically developing controls, using a 16-channel acquisition system. The signals were filtered using a fourth-order Butterworth band-pass filter, transformed into PCA or MS-PCA representations, segmented into 4 s windows with 50% overlap, and evaluated using subject-wise 5-fold cross-validation to reduce subject-identity leakage. The results showed that MS-PCA produced higher descriptive performance than PCA in both temporal architectures, with the strongest descriptive result obtained by the MS-PCA + TCN-SA scheme, which achieved a mean accuracy of 97.96 ± 2.37% and balanced precision, recall, F1-score, and specificity. However, the inferential comparison between PCA and MS-PCA did not reach statistical significance at the 0.05 level, and the cohort size was limited to 10 participants. Therefore, these findings should be interpreted as preliminary descriptive evidence within the present cohort rather than evidence of diagnostic readiness or robust clinical applicability. Larger, independent, and demographically diverse EEG datasets with richer clinical characterization are required to confirm the observed trend and evaluate the generalizability of the proposed framework.**

Keywords **Autism Spectrum Disorder; Electroencephalography; Multiscale Principal Component Analysis; Temporal Convolutional Network; Bidirectional Long Short-Term Memory.**

I. Introduction

Autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental condition associated with impairments in social communication and restricted or repetitive patterns of behavior [1], [2]. Early identification remains difficult because diagnosis still depends heavily on behavioral observation and expert judgment, which can be time-consuming and influenced by clinical heterogeneity [2], [3]. These limitations have encouraged the search for more objective biomarkers to complement conventional assessment [3], [4]. Electroencephalography (EEG) is

Manuscript received 7 May 2026; Revised 5 July 2026; Accepted 15 July 2026; Available online 23 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1735

a promising modality for ASD-related biomarker research because it is noninvasive, relatively low-cost, and provides high temporal resolution for capturing neural dynamics [4], [5]. Previous studies have reported alterations in EEG power, coherence, and functional connectivity in ASD across multiple temporal and frequency scales [6], [7]. However, scalp EEG signals are highly variable, nonstationary, and susceptible to diverse artifacts, making rigorous data quality control and preprocessing essential before modeling [8], [9], [10]. This challenge becomes more critical in studies with limited sample sizes, where instability of learned EEG features and model-selection choices may reduce robustness and generalizability [11], [12].

Feature extraction is, therefore, an important step in constructing compact and informative EEG representations. Principal component analysis (PCA) has been widely used to reduce inter-channel redundancy and summarize complex time-frequency or multichannel EEG structure by projecting data onto orthogonal components ordered by explained variance [11], [13], [14]. However, conventional PCA provides a single-scale linear projection and may not sufficiently preserve localized temporal patterns in nonstationary EEG signals. Multiscale principal component analysis (MS-PCA) offers a potential alternative because it can represent complementary signal structures across multiple temporal scales [11], [15]. More broadly, recent EEG-based machine-learning studies have emphasized the importance of feature engineering and model selection for biomarker-oriented classification pipelines [16]. In autism-related EEG research, structured feature representations, including network-topology and graph-based features, have also shown potential for improving discriminative learning from complex neural signals [17], [18], [19].

After feature extraction, temporal modeling is needed to learn dependencies within EEG windows. Recurrent architectures, including bidirectional long short-term memory (BiLSTM) networks, have been used for EEG sequence learning and classification because they can model temporal dependencies in sequential data [20], [21]. Related EEG studies have also shown that recurrent models can outperform more conventional baselines in sequence-oriented settings [22]. In parallel, temporal convolutional networks (TCNs) can enlarge the effective receptive field through dilated convolutions while maintaining efficient optimization on long sequences [23]. When combined with self-attention, temporal models may further emphasize informative temporal patterns within EEG sequences [24]. Recent autism-focused work also indicates that deep architectures combining spatial and temporal representations can support EEG-based ASD classification research [18], [25].

Despite these advances, EEG-based ASD classification studies still face two important methodological challenges. First, although PCA and deep learning models have been widely explored, the interaction between multiscale feature representation and temporal learning architecture remains less frequently evaluated under the same subject-wise validation protocol [11], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25]. Second, EEG classification commonly uses overlapping windows or EEG-derived segments, and segment-level splitting may overestimate performance when data from the same participant appear in both training and testing partitions. Therefore, subject-wise fold assignment should be performed before feature fitting, normalization, windowing, model training, and testing to reduce subject-identity leakage [9], [26].

To address this focused gap, this study compares PCA and MS-PCA feature representations, paired with RNN-BiLSTM and TCN-SA models, under a leakage-aware, subject-wise 5-fold cross-validation protocol. In this framework, PCA/MS-PCA fitting and normalization parameters are estimated only from training subjects and then applied unchanged to held-out subjects. The aim is not to claim diagnostic readiness, but to examine whether a multiscale, component-based representation is associated with descriptive differences in temporal learning behavior and classification performance within a preliminary EEG-based ASD cohort.

The main contributions of this study are as follows. First, this study proposes a leakage-aware comparative EEG classification framework that directly evaluates single-scale PCA and MS-PCA feature representations for ASD classification. Second, this study investigates the interaction between feature representation and temporal learning architecture by pairing PCA and MS-PCA with both RNN-BiLSTM and TCN-SA models. Third, this study evaluates the MS-PCA + TCN-SA scheme as a multiscale component-based temporal learning pipeline for modeling complementary EEG structures and within-window temporal dependencies. Fourth, this study assesses the proposed schemes using subject-wise 5-fold cross-validation, window-level and subject-level summaries, multiple classification metrics, and exact Wilcoxon signed-rank testing with rank-biserial correlation.

The remainder of this paper is organized as follows. Section II describes the participants, EEG acquisition protocol, preprocessing, feature extraction, temporal deep learning models, and statistical analysis. Section III presents the training dynamics, classification performance, and statistical comparison results. Section IV discusses the interpretation of the findings, comparison with related studies, limitations, and

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

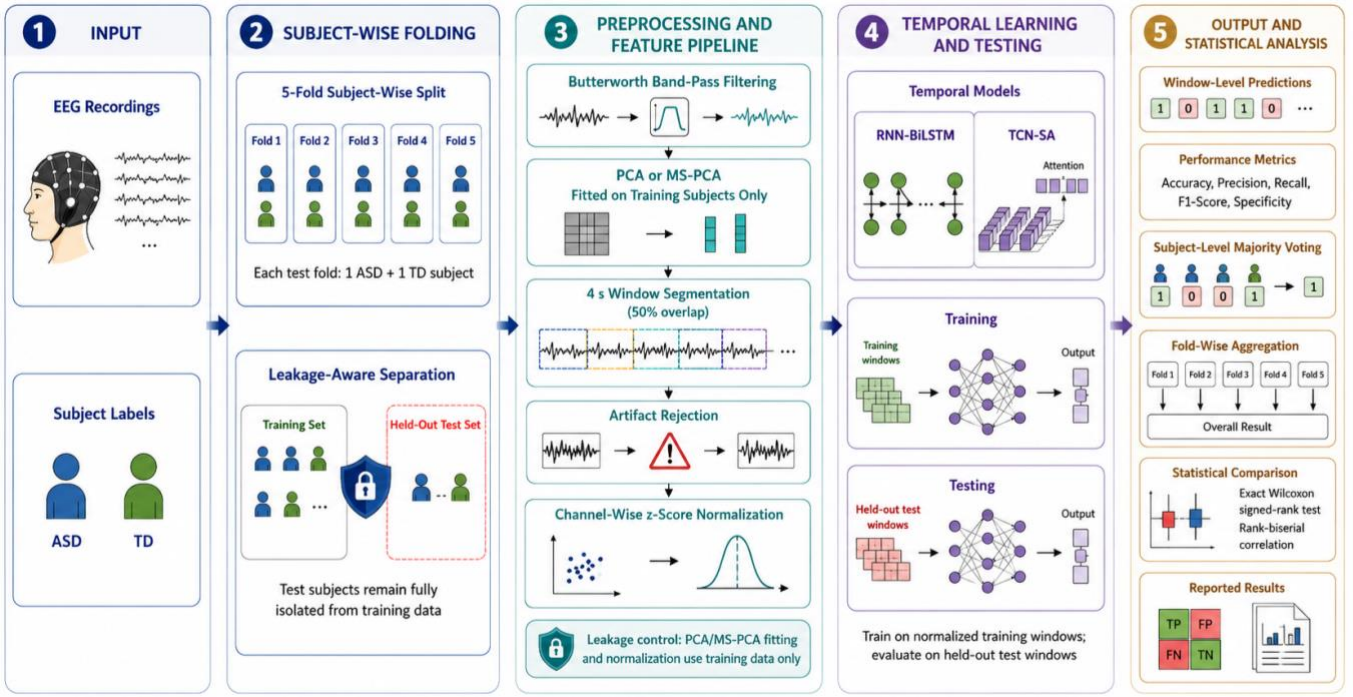

Fig. 1. Overall workflow of the proposed leakage-aware EEG classification pipeline for ASD. The figure summarizes the complete pipeline from EEG recordings and subject labels to subject-wise fold assignment, preprocessing and feature extraction, temporal learning and testing, output evaluation, and statistical analysis. Subject-wise fold assignment is performed before feature fitting, windowing, normalization, model training, and model evaluation to reduce subject-identity leakage.

implications. Section V concludes the study and outlines future research directions.

II. Method

This study implemented a comparative deep learning pipeline for classifying autism spectrum disorder and typically developing controls using multichannel electroencephalographic signals. The workflow consisted of participant recruitment, EEG data acquisition, subject-wise fold assignment, band-pass filtering, feature extraction, windowing, normalization, temporal classification, performance evaluation, and statistical analysis. Two feature representations were compared, namely principal component analysis and multiscale principal component analysis. Each representation was evaluated using two temporal learning models, namely a recurrent neural network with bidirectional long short-term memory and a temporal convolutional network with self-attention. The overall workflow of the proposed study is shown in Fig. 1.

Fig. 1 highlights that subject-wise fold assignment was performed before window generation and feature fitting, which is important because overlapping windows from the same participant could otherwise introduce subject-identity leakage.

To further clarify the complete leakage-aware implementation procedure, the proposed pipeline is summarized in Algorithm 1. The algorithm emphasizes that subject-wise fold assignment was performed before feature fitting, windowing, normalization, model training, and testing. It also shows that PCA/MS-PCA parameters and normalization statistics were estimated only from the training subjects within each fold and then applied unchanged to the held-out subjects.

Algorithm 1. Leakage-aware PCA/MS-PCA temporal learning pipeline for EEG-based ASD classification

Input: EEG recordings from all participants; subject labels; feature representations $R = \{\text{PCA}, \text{MS-PCA}\}$; temporal models $M = \{\text{RNN-BiLSTM}, \text{TCN-SA}\}$; number of folds $K = 5$.

Output: Window-level classification metrics, subject-level majority-voting results, and statistical comparison between PCA and MS-PCA schemes.

1. Assign participants into five subject-wise folds, with each test fold containing one ASD participant and one typically developing participant.
2. For each fold $k = 1, \ldots, 5$:
   2.1 Define the training subjects and held-out test subjects.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

2.2 Keep all data from held-out test subjects completely separate from the training partition.
2.3 Apply Butterworth band-pass filtering to the EEG recordings according to the predefined preprocessing setting.
3. For each feature representation $r \in R$:
3.1 Fit PCA or MS-PCA parameters using only the training subjects in the current fold.
3.2 Apply the fitted PCA/MS-PCA transformation unchanged to both training subjects and held-out test subjects.
3.3 Segment the transformed EEG representations into 4 s windows with 50% overlap.
3.4 Reject windows with amplitudes exceeding the predefined artifact threshold.
3.5 Estimate channel-wise z-score normalization parameters using only the training windows.
3.6 Apply the training-derived normalization parameters unchanged to the held-out test windows.
4. For each temporal model $m \in M$:
4.1 Train the model using only the normalized training windows.
4.2 Test the trained model on the normalized held-out test windows.
4.3 Compute window-level predictions and aggregate TP, TN, FP, and FN.
4.4 Calculate accuracy, precision, recall, F1-score, and specificity for the current fold.
4.5 Aggregate window-level predictions within each held-out subject using majority voting to obtain subject-level predictions.
5. Repeat Steps 2–4 until every participant appears once in the held-out test set.
6. Aggregate the fold-wise performance metrics across the five folds.
7. Compare PCA and MS-PCA within each temporal architecture using the exact Wilcoxon signed-rank test based on fold-wise accuracy values.
8. Calculate rank-biserial correlation to quantify the effect magnitude of the paired comparison.
9. Report window-level metrics, subject-level majority-voting summaries, confusion matrices, and statistical comparison results.

A. Participants and Data Acquisition

The study included 10 participants, consisting of 5 individuals with autism spectrum disorder and 5 typically developing (TD) controls. All participants were between 15 and 25 years of age to reduce age-related variability in the electroencephalographic recordings. The ASD group consisted of three male and two female participants, whereas the TD control group consisted of five male participants. The diagnostic labels and acquisition protocol were based on the previously reported EEG dataset described in [10]. ASD diagnoses were confirmed using clinical records at the treating facility, whereas the control participants had no reported history of neurological or psychiatric disorders. Detailed individualized clinical metadata, including standardized ASD severity scores, cognitive level, medication status, and comorbidity profiles, were not available as complete structured variables for all participants in the present analysis. Therefore, these variables were not used as covariates in model training or statistical comparisons. This limitation should be considered when interpreting the findings, because the present dataset may not fully capture the broader clinical heterogeneity of ASD. Written informed consent was obtained from all participants or their legal guardians prior to data collection. The study protocol was approved by the Health Research Ethics Committee of the Faculty of Medicine, Universitas Syiah Kuala, Indonesia, with approval number 117/EA/FK/2024. Given the limited cohort size, this study was designed as a preliminary methodological evaluation of PCA/MS-PCA-based temporal learning rather than as a population-level validation of EEG-based ASD classification.

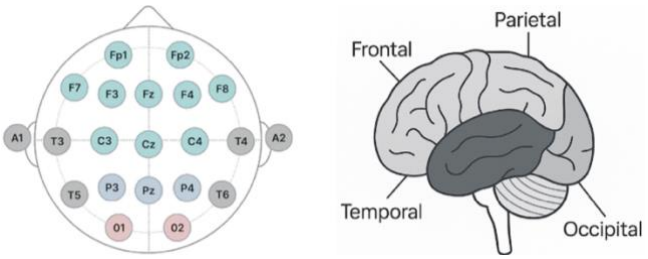

Fig. 2. Electroencephalographic 10–20 electrode montage and cortical-lobe anatomy. The figure illustrates the 16-channel electrode distribution used in this study and shows the cortical regions represented by the frontal, temporal, central, parietal, and occipital recording sites.

EEG data were recorded under a resting-state eyes-open condition using a 16-channel OpenBCI Cyton-based acquisition system [27]. Electrodes were positioned according to the international 10–20 system at Fp1, Fp2, F7, F3, F4, F8, T7, C3, C4, T8, P7, P3, P4, P8, O1, and O2. The electrode layout and cortical-lobe illustration are presented in Fig. 2. Recordings were conducted in the morning in a quiet and controlled environment to minimize environmental variability. For each participant, the recorded EEG signal lasted approximately 15 min. The EEG acquisition procedure, from electrode-cap placement to raw data storage, is presented schematically in Fig. 3. Raw electroencephalographic signals were saved in text format together with session notes and quality information [27].

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

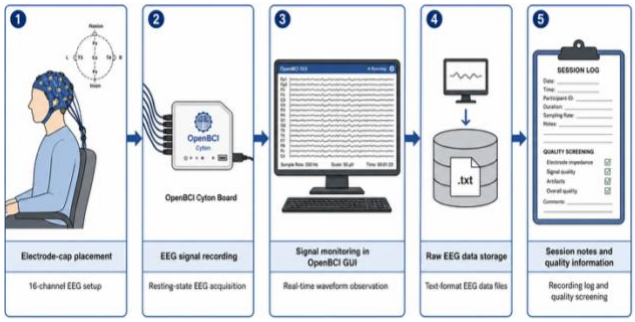

Fig. 3. Schematic EEG data acquisition workflow. The diagram shows electrode-cap placement, OpenBCI Cyton-based signal recording, OpenBCI software monitoring, and raw EEG data storage for subsequent preprocessing. No identifiable participant photograph is shown.

Fig. 2 clarifies the spatial coverage of the 16-channel montage used in this study, whereas Fig. 3 illustrates the practical acquisition procedure from cap placement to raw EEG storage, supporting reproducibility of the recording protocol.

B. Leakage-Aware Cross-Validation Protocol

To minimize the risk of subject-identity leakage from overlapping EEG windows, the cross-validation procedure was defined at the participant level before any feature fitting, normalization, windowing, model training, or evaluation was performed. The 10 participants were divided into five subject-wise folds, with each held-out test fold containing one ASD participant and one typically developing participant. In each fold, all data from the held-out participants were kept completely separate from the training partition. Consequently, all overlapping windows for a given participant remained exclusively within a single partition and were never split between training and testing.

PCA and MS-PCA fitting were performed using only the training subjects within each fold. The resulting projection matrices, retained components, and multiscale transformation parameters were then applied unchanged to the corresponding held-out test subjects. Similarly, channel-wise z-score normalization parameters were estimated only from the training partition and then applied to the held-out data without recalculation. The held-out test folds were not used for feature fitting, normalization-parameter estimation, hyperparameter selection, early stopping, or model selection, but only for final performance evaluation. This fold-wise procedure was repeated until each participant appeared once in the held-out test set.

C. Butterworth Band-Pass Filtering

Within each subject-wise fold, the continuous raw electroencephalographic signals were first processed using a fourth-order Butterworth band-pass filter with cut-off frequencies of 0.5–40 Hz. The filter was applied in a forward–reverse manner to obtain zero-phase filtering while suppressing low-frequency drift and high-frequency noise [8], [10]. As illustrated in Fig. 4, the filtering stage was intended to reduce baseline drift and high-frequency noise while retaining the 0.5–40 Hz EEG range used for feature extraction. After Butterworth band-pass filtering, the filtered signals were forwarded to two alternative feature-generation branches, namely principal component analysis and multiscale principal component analysis, following the overall workflow shown in Fig. 1 [15], [28]. Fig. 5 illustrates the PCA and MS-PCA feature-generation workflows, including training-fold fitting and application of the fitted transformations to held-out subjects.

EEG signal quality was controlled during acquisition by conducting recordings in a quiet environment, checking electrode contact before data collection, and documenting session-level quality information together with the raw EEG files. After acquisition, the continuous EEG recordings were screened to identify channels or segments with abnormally large amplitudes, unstable contact, or nonphysiological signal patterns. The main preprocessing steps consisted of zero-phase Butterworth band-pass filtering and amplitude-based window rejection. Windows with amplitudes exceeding ±100 μV were discarded to reduce the influence of large residual artifacts, including abrupt movement-related artifacts. Because no dedicated electrooculography or electromyography channels were recorded, eye-movement and muscle-related artifacts were not removed using independent component analysis or regression-based correction. This limitation is acknowledged because residual ocular, muscular, or bad-channel artifacts may still affect the extracted EEG features. Numerical electrode impedance values were not available for all recordings; therefore, electrode-contact stability and session-level signal quality were checked and documented before data collection.

D. Principal Component Analysis

Principal component analysis was used as the first feature-generation branch to reduce inter-channel redundancy and obtain an orthogonal representation for downstream temporal modeling [13], [14]. The conceptual workflow of PCA and its distinction from MS-PCA are illustrated in Fig. 5(a). The equations in the feature-extraction subsections are used to define the exact transformations applied in the implementation. Specifically, they describe how the filtered multichannel EEG matrix was transformed into either a single-scale PCA representation or a multiscale PCA representation before windowing and temporal classification. All feature-fitting operations described by these equations were performed only on the training subjects within each fold and then applied unchanged to the corresponding held-out subjects.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

The PCA formulation in Eqs. (1)–(4) follows the standard covariance-based PCA procedure commonly used for dimensionality reduction, redundancy reduction, and orthogonal EEG feature representation [13], [14]. For each fold, the filtered multichannel EEG signal was arranged as a time-by-channel matrix (Eq. (1) [13], [14]):

$$\mathbf{X} \in \mathbb{R}^{T \times C} \quad (1)$$

where T denotes the number of time samples and C denotes the number of EEG channels. Before covariance estimation, $\mathbf{X}$ was mean-centered channel-wise. The covariance matrix was computed as Eq. (2) [13], [14]:

$$\mathbf{\Sigma} = \frac{1}{T-1}\mathbf{X}^{\top}\mathbf{X} \quad (2)$$

where $\mathbf{\Sigma} \in \mathbb{R}^{C \times C}$ denotes the channel covariance matrix and $(\cdot)^{\top}$ denotes matrix transpose. Eigenvalue decomposition was then performed as Eq. (3):

$$\mathbf{\Sigma V} = \mathbf{V\Lambda} \quad (3)$$

where $\mathbf{V}$ contains the eigenvectors and $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues sorted in descending order. The filtered signal was then projected onto the retained principal components as Eq. (4):

$$\mathbf{S}_{\mathrm{PCA}} = \mathbf{X}\mathbf{V}_k \quad (4)$$

where $\mathbf{S}_{\mathrm{PCA}} \in \mathbb{R}^{T \times k}$ denotes the PCA score matrix and $\mathbf{V}_k$ contains the first k eigenvectors. In this study, $k = 16$ principal components were retained. The final PCA representation was expressed as $\mathbf{Z}_{\mathrm{PCA}} = \mathbf{S}_{\mathrm{PCA}}^{\top} \in \mathbb{R}^{16 \times T}$, enabling direct comparison with the MS-PCA branch. In the implementation, Eqs. (1)–(4) correspond to the fold-wise PCA procedure: arranging the filtered EEG signal into a time-by-channel matrix, estimating the training-fold covariance matrix, computing the eigenvector basis, and projecting both training and held-out data onto the retained 16 principal components. The PCA projection basis was estimated only from the training subjects within each fold and then applied unchanged to the corresponding held-out test subjects; no held-out test data were used in covariance estimation, eigenvector computation, or component selection [9].

E. Multiscale Principal Component Analysis

Multiscale principal component analysis was used as the second feature-generation branch to capture the structure of the EEG signal across multiple temporal scales [15]. The conceptual workflow of the MS-PCA branch is illustrated in Fig. 5(b). In this approach, the Butterworth-filtered EEG signal $\mathbf{X}$ was first smoothed using Gaussian kernels with different scale parameters, and PCA was then applied independently at each scale. The MS-PCA formulation in Eqs. (5)–(10) follows the standard multiscale PCA concept, in which smoothed representations are generated at multiple temporal scales before PCA is applied at each scale [15]. The Gaussian kernel was defined as Eq. (5):

$$g_{\sigma}(\tau) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\tau^2}{2\sigma^2}\right) \quad (5)$$

where σ denotes the smoothing scale and τ denotes the temporal lag. The smoothed signal at scale σ was obtained as Eq. (6):

$$\mathbf{X}^{(\sigma)} = \mathbf{X} * g_{\sigma} \quad (6)$$

where $*$ denotes one-dimensional convolution along the time axis. In this study, four smoothing scales were

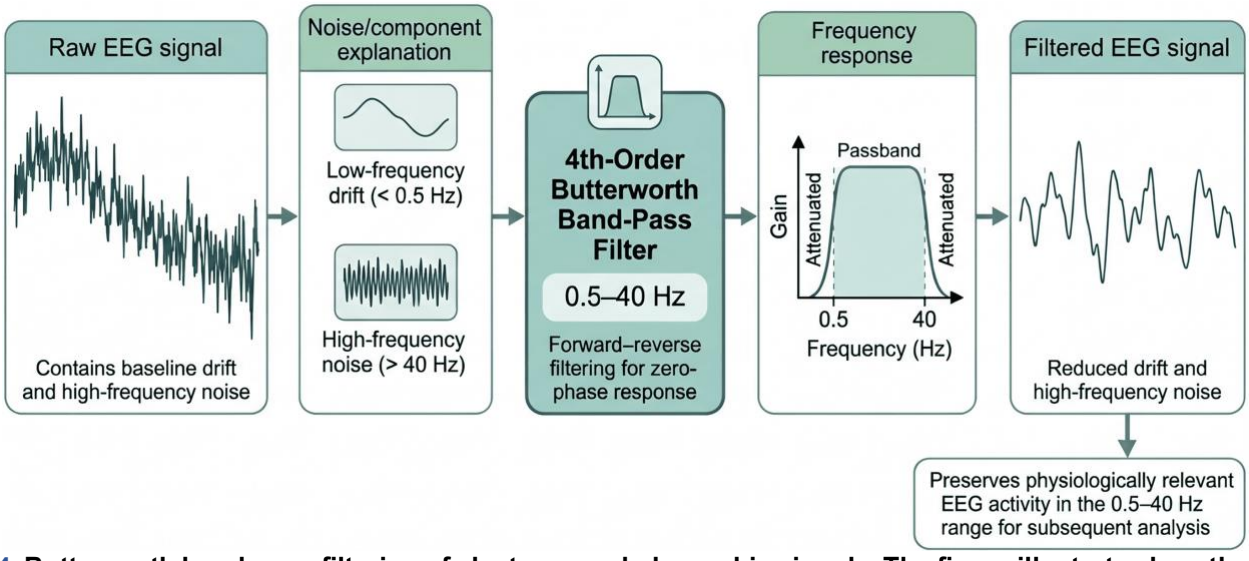

Fig. 4. Butterworth band-pass filtering of electroencephalographic signals. The figure illustrates how the 0.5–40 Hz fourth-order Butterworth filter suppresses low-frequency drift and high-frequency noise while preserving the physiologically relevant EEG frequency range used for subsequent feature extraction.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

used, namely $\sigma \in \{1,4,8,16\}$. For each scale, a separate covariance matrix was computed as Eq. (7):

$$\mathbf{\Sigma}^{(\sigma)} = \frac{1}{T-1}\left(\mathbf{X}^{(\sigma)}\right)^{\top}\mathbf{X}^{(\sigma)} \quad (7)$$

where $\mathbf{\Sigma}^{(\sigma)}$ denotes the covariance matrix at the smoothing scale σ. Eigenvalue decomposition was then performed as Eq. (8):

$$\mathbf{\Sigma}^{(\sigma)}\mathbf{V}^{(\sigma)} = \mathbf{V}^{(\sigma)}\mathbf{\Lambda}^{(\sigma)} \quad (8)$$

where $\mathbf{V}^{(\sigma)}$ contains the eigenvectors and $\mathbf{\Lambda}^{(\sigma)}$ is the diagonal matrix of eigenvalues for scale σ. The score matrix at each scale was obtained by projecting the smoothed signal onto the retained principal components as Eq. (9):

$$\mathbf{S}^{(\sigma)} = \mathbf{X}^{(\sigma)}\mathbf{V}_{k_\sigma}^{(\sigma)} \quad (9)$$

where $\mathbf{V}_{k_\sigma}^{(\sigma)}$ contains the first k_σ eigenvectors retained at scale σ, and $\mathbf{S}^{(\sigma)} \in \mathbb{R}^{T\times k_\sigma}$ denotes the score matrix at that scale. In this study, $k_\sigma = 4$ principal components were retained for each smoothing scale. The score matrices from the four scales were concatenated along the component dimension as Eq. (10):

$$\mathbf{S}_{\mathrm{MSPCA}} = [\mathbf{S}^{(1)} \parallel \mathbf{S}^{(4)} \parallel \mathbf{S}^{(8)} \parallel \mathbf{S}^{(16)}] \in \mathbb{R}^{T\times 16} \quad (10)$$

where $\parallel$ denotes concatenation along the feature or component dimension. The final MS-PCA representation was expressed as $\mathbf{Z}_{\mathrm{MSPCA}} = \mathbf{S}_{\mathrm{MSPCA}}^{\top} \in \mathbb{R}^{16\times T}$. Thus, each filtered EEG signal was represented by 16 pseudo-channels, corresponding to four retained components from each of four temporal scales [15].

In the implementation, Eqs. (5)–(10) define the MS-PCA branch: the filtered EEG signal was smoothed at four temporal scales, PCA was applied separately at each scale using only the training subjects, four components were retained per scale, and the resulting score matrices were concatenated to obtain a $16 \times T$ multiscale pseudo-channel representation. As in the PCA branch, all MS-PCA fitting steps, including covariance estimation, eigenvector computation, and retained-component projection at each temporal scale, were performed using only the training subjects within each fold and then applied unchanged to the held-out test subjects to maintain leakage-free evaluation [9]. Fig. 5 provides the key methodological distinction between PCA and MS-PCA: PCA extracts components from a single filtered representation, whereas MS-PCA extracts and concatenates components from multiple smoothed temporal scales.

F. Windowing and Normalization

After feature generation using either principal component analysis or multiscale principal component analysis, the resulting signal representations were segmented into 4-s windows with 50% overlap, following the sequence illustrated in Fig. 1. Let $\mathbf{Z} \in \mathbb{R}^{C\times T}$ denote the feature representation obtained from either branch. The n-th window was defined as Eq. (11):

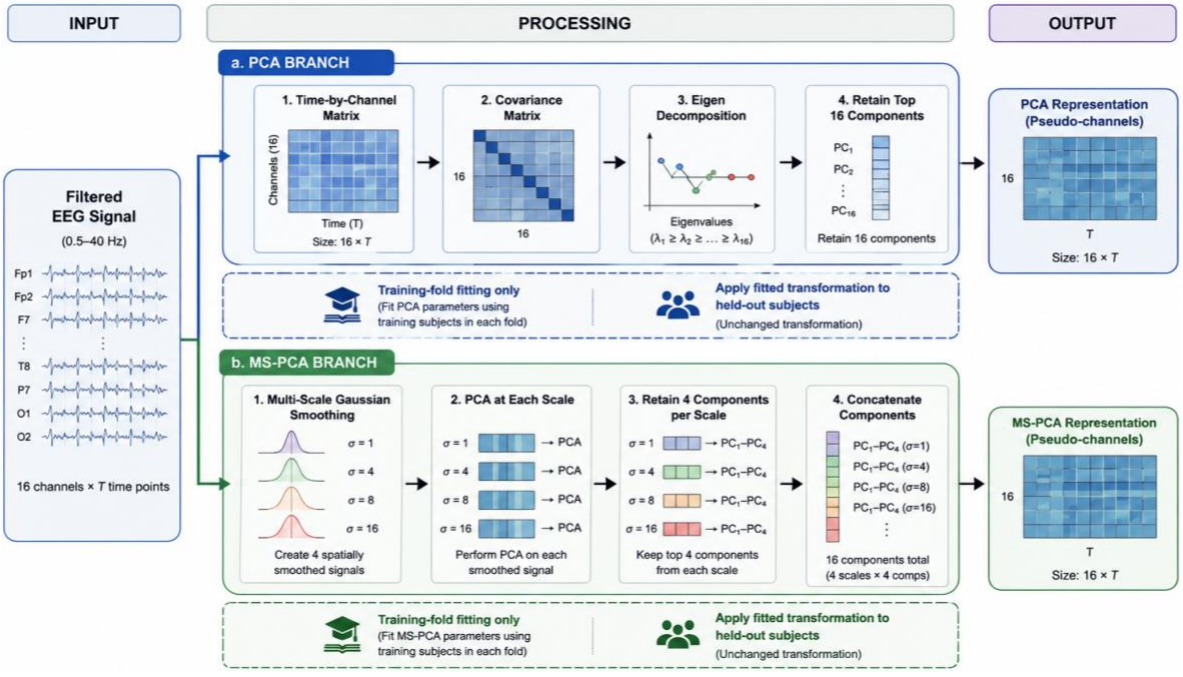

Fig. 5. PCA and MS-PCA feature-generation workflows for EEG representation. The diagram summarizes the single-scale PCA and multiscale PCA branches, including training-fold fitting and application of the fitted transformations to held-out subjects.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

$$\mathbf{W}_n = \mathbf{Z}[:, nL : nL + W - 1] \quad (11)$$

where n denotes the window index, $W = 1000$ denotes the window length, $L = 500$ denotes the stride corresponding to 50% overlap at the sampling rate of 250 Hz, and ":" denotes selection of all feature channels. Therefore, each segment contained approximately 1000 samples per channel. Windows with amplitudes exceeding $\pm 100\ \mu V$ were discarded to reduce the influence of large residual artifacts, following standard EEG quality-control practice [8]. This stage produced the final input segments used for classification. Because fold assignment had been completed at the participant level before window generation, overlapping windows from the same participant were assigned only to the partition of that participant and could not appear simultaneously in the training and held-out test sets. Following window generation, channel-wise z-score normalization was applied at the fold level. The mean and standard deviation were estimated only from the training subjects in each fold and were then applied unchanged to the held-out test subjects in order to prevent information leakage [9], [28]. The normalized window matrix was expressed as Eq. (12):

$$\widetilde{\mathbf{W}}_n(t,c) = \frac{\mathbf{W}_n(t,c) - \mu_c^{(\text{train})}}{\sigma_c^{(\text{train})}} \quad (12)$$

where $\widehat{\mathbf{W}}_n(t,c)$ denotes the normalized value at time index t and channel c in the n-th window, $\mathbf{W}_n(t,c)$ denotes the original window amplitude, and $\mu_c^{(\text{train})}$ and $\sigma_c^{(\text{train})}$ denote the mean and standard deviation estimated only from the training data of the corresponding fold. In the implementation, Eqs. (11)–(12) define the segmentation and normalization steps used before model input. The window length and stride were fixed at 1000 and 500 samples, respectively, and the normalization parameters were estimated only from the training partition of each fold to prevent information leakage. The normalized windowed representations obtained from the principal component analysis and multiscale principal component analysis branches were then used as inputs to the RNN-BiLSTM and TCN-SA models.

G. RNN-BiLSTM

The first temporal learning model combined a recurrent neural network layer with a bidirectional long short-term memory layer to capture sequential dependencies in both temporal directions [20], [21]. Each input segment consisted of 1000 time steps with 16 features per time step. Fig. 6 illustrates the hidden-state flow of the recurrent unit, whereas Fig. 7 shows how the BiLSTM layer combines forward and backward temporal contexts for EEG sequence classification. The following equations specify the temporal operations represented by the RNN-BiLSTM model used in this study. They are included to clarify how sequential EEG features were propagated through recurrent hidden states, processed in forward and backward temporal directions, pooled, and mapped to a binary output. In the implementation, these operations were realized using PyTorch recurrent and LSTM layers with the hyperparameters reported in Table 1.

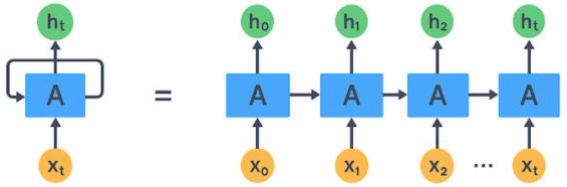

Fig. 6. Hidden-state flow in an RNN for EEG-based classification. The figure illustrates how recurrent hidden states transfer temporal information across successive EEG time steps before downstream bidirectional sequence modelling.

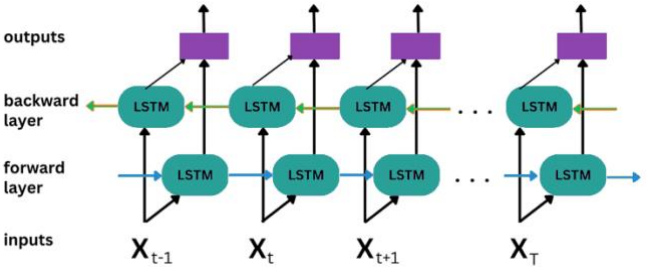

Fig. 7. Bidirectional LSTM architecture. The figure shows how forward and backward LSTM layers process EEG sequences in both temporal directions to capture contextual dependencies before classification.

The recurrent and LSTM gate equations in Eqs. (13)–(24) follow the standard RNN and LSTM formulations used for sequence learning [20], [21]. Let $\mathbf{x}_t \in \mathbb{R}^{16}$ denote the feature vector at time step t. The hidden state of the recurrent neural network layer was updated as Eq. (13):

$$\mathbf{h}_t = \phi(\mathbf{W}_{xh}\mathbf{x}_t + \mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{b}_h) \quad (13)$$

where $\mathbf{h}_t$ denotes the hidden state at time step t, $\phi(\cdot)$ is a nonlinear activation function, and $\mathbf{W}_{xh}$, $\mathbf{W}_{hh}$, and $\mathbf{b}_h$ are the learnable parameters.

The bidirectional long short-term memory layer was used to model contextual dependencies in both forward and backward directions. For each direction, the input, forget, and output gates were defined as Eq. (14), Eq. (15), Eq. (16), and Eq. (17):

$$\mathbf{i}_t = \sigma(\mathbf{W}_{xi}\mathbf{x}_t + \mathbf{W}_{hi}\mathbf{h}_{t-1} + \mathbf{b}_i) \quad (14)$$

$$\mathbf{f}_t = \sigma\left(\mathbf{W}_{xf}\mathbf{x}_t + \mathbf{W}_{hf}\mathbf{h}_{t-1} + \mathbf{b}_f\right) \quad (15)$$

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

$$\mathbf{o}_t = \sigma(\mathbf{W}_{xo}\mathbf{x}_t + \mathbf{W}_{ho}\mathbf{h}_{t-1} + \mathbf{b}_o) \quad (16)$$

$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_{xg}\mathbf{x}_t + \mathbf{W}_{hg}\mathbf{h}_{t-1} + \mathbf{b}_g) \quad (17)$$

where $\mathbf{i}_t$, $\mathbf{f}_t$, and $\mathbf{o}_t$ denote the input, forget, and output gates, respectively; $\tilde{\mathbf{c}}_t$ denotes the candidate cell state; $\sigma(\cdot)$ denotes the sigmoid activation function; $\tanh(\cdot)$ denotes the hyperbolic tangent function; and $\mathbf{W}$ and $\mathbf{b}$ terms denote learnable gate-specific weights and biases. The cell state and hidden state were then updated as Eq. (18) and Eq. (19):

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t \quad (18)$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t) \quad (19)$$

where $\odot$ denotes element-wise multiplication. For the bidirectional structure, the forward and backward hidden representations were computed as Eq. (20) and Eq. (21):

$$\vec{\mathbf{h}}_t = \mathrm{LSTM}_{\mathrm{fwd}}(\mathbf{x}_{1:t}) \quad (20)$$

$$\overleftarrow{\mathbf{h}}_t = \mathrm{LSTM}_{\mathrm{bwd}}(\mathbf{x}_{T:t}) \quad (21)$$

where $\mathbf{x}_{1:t}$ denotes the forward input sequence from time step 1 to t, whereas $\mathbf{x}_{T:t}$ denotes the backward input sequence from time step T to t. The bidirectional hidden representation was then formed by concatenating the forward and backward outputs: as Eq. (22):

$$\mathbf{h}_t^{\mathrm{bi}} = [\vec{\mathbf{h}}_t \; ; \; \overleftarrow{\mathbf{h}}_t] \quad (22)$$

where $[\cdot\,;\,\cdot]$ denotes vector concatenation. The sequence-level representation was then pooled over time and mapped to a binary output logit as Eq. (23) and Eq. (24):

$$z = \mathbf{W}_o \,\mathrm{Pool}\left(\{\mathbf{h}_t^{\mathrm{bi}}\}_{t=1}^{T}\right) + b_o \quad (23)$$

$$\hat{y} = \sigma(z) \quad (24)$$

where $\mathrm{Pool}(\cdot)$ denotes temporal pooling over the sequence dimension, $\mathbf{W}_o$ and b_o denote the learnable output-layer weight and bias, z denotes the output logit, and $\hat{y}$ denotes the predicted probability for the positive class. For the principal component analysis representation, the model used a recurrent neural network layer with 128 hidden units followed by a bidirectional long short-term memory layer with 128 units per direction stacked in four layers. For the multiscale principal component analysis representation, a lighter configuration was used, consisting of a recurrent neural network layer with 32 hidden units and a bidirectional long short-term memory layer with 32 units per direction stacked in two layers. In both configurations, the recurrent outputs were followed by two fully connected layers and a single output logit for binary classification. Dropout was applied after the recurrent block and before the dense classifier to reduce overfitting [20]. Training was conducted using the Adamax optimizer with a learning rate of 1×10^{-4}, a batch size of 64, and 50 epochs. Binary cross-entropy with logits loss was used as the optimization objective. The configurations are summarized in Table 1.

Table 1. RNN-BiLSTM configurations for PCA and MS-PCA.

Component	Hyperparameters	PCA	MS-PCA
RNN layer	Hidden units	128	32
BiLSTM layer	Hidden units per direction	128	32
	Number of layers	4	2
FC layer	Number of FC	2	2
Output layer	Loss	BCEWithLogitsLoss	BCEWithLogitsLoss
Training	Batch size	64	64
	Epochs	50	50
	Optimizer	Adamax	Adamax
	Learning rate	0.0001	0.0001

H. TCN-SA

The second temporal learning model was a temporal convolutional network with self-attention (TCN-SA), which combines dilated causal one-dimensional convolutions with an attention-based reweighting mechanism [23], [24]. The dilated convolutional blocks were used to increase the effective receptive field while preserving temporal resolution, whereas the self-attention layer was used to emphasize informative temporal regions within each EEG segment. Fig. 8 illustrates the residual TCN-SA structure used in this study, including dilated temporal convolutions to capture a broader temporal context and self-attention to emphasize informative EEG time regions.

The following equations define the main operations implemented in the TCN-SA branch. The dilated convolution equations describe how the model expands the temporal receptive field within each EEG window, whereas the self-attention equation describes

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

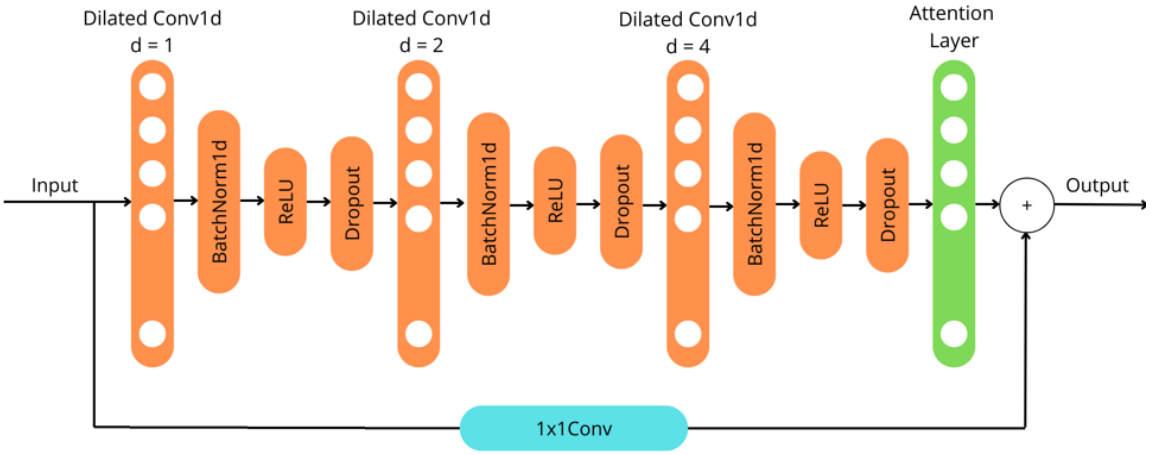

Fig. 8. TCN with self-attention for electroencephalographic classification. The figure illustrates how dilated temporal convolutional blocks extract temporal patterns at increasing receptive fields, while the self-attention layer emphasizes informative time regions before binary classification.

how informative temporal regions are reweighted before binary classification. These operations were implemented using the PyTorch model configuration summarized in Table 2. The dilated convolution and receptive-field formulations in Eqs. (25)–(26) follow the standard temporal convolutional network formulation [23]. Let $\mathbf{x}_t \in \mathbb{R}^{16}$ denote the input feature vector at time step t. A dilated causal convolution with a kernel size k and dilation factor d was expressed as Eq. (25):

$$\mathbf{y}_t = \sum_{i=0}^{k-1} \mathbf{W}_i \, \mathbf{x}_{t-di} + \mathbf{b} \tag{25}$$

where $\mathbf{y}_t$ denotes the convolution output at time step t, i denotes the kernel lag index, k denotes the kernel size, d denotes the dilation factor, $\mathbf{W}_i$ denotes the convolution kernel at lag i, and $\mathbf{b}$ denotes the bias term.

The effective receptive field of a stack of L dilated convolutional layers was defined as Eq. (26):

$$R = 1 + (k-1)\sum_{\ell=1}^{L} d_\ell \tag{26}$$

where d_ℓ denotes the dilation factor in the ℓ-th layer. This formulation allows the model to capture a wider temporal context without reducing the sequence resolution. The self-attention operation in Eq. (27) follows the standard scaled dot-product attention formulation [24]. Let $\mathbf{H} \in \mathbb{R}^{T\times D}$ denote the feature sequence produced by the temporal convolutional blocks. The self-attention mechanism was then computed as:

$$\mathrm{Attn}(\mathbf{H}) = \mathrm{softmax}\left(\frac{(\mathbf{W}_Q)(\mathbf{H}\mathbf{W}_K)^{\top}}{\sqrt{d_k}}\right)(\mathbf{H}\mathbf{W}_V) \tag{27}$$

Table 2. TCN-SA configurations for PCA and MS-PCA.

Component	Hyperparameters	PCA	MS-PCA
TCN blocks	Channels	32–64–64	16–32–32
	Kernel size	7	5
	Dilations	1–2–4	1–2–4
Self-attention	Heads	4	4
	Embedding dimension	64	32
FC layer	Units	1	1
Output layer	Loss	BCEWithLogitsLoss	BCEWithLogitsLoss
Training	Batch size	64	64
	Epochs	50	50
	Optimizer	Adamax	Adamax
	Learning rate	0.0001	0.0001

Note: The channel notation 32–64–64 and 16–32–32 denotes the output channel widths of the first, second, and third TCN blocks for the PCA and MS-PCA branches, respectively. The dilation notation 1–2–4 denotes the dilation factors applied sequentially to the first, second, and third TCN blocks.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

where $\mathbf{W}_Q$, $\mathbf{W}_K$, and $\mathbf{W}_V$ denote the projection matrices for query, key, and value, respectively, and d_k denotes the key dimension. The attended feature sequence was pooled over time and projected to a single output logit as Eq. (28) and Eq. (29):

$$z = \mathbf{W}_o \, \mathrm{Pool}\big(\mathrm{Attn}(\mathbf{H})\big) + b_o \tag{28}$$

$$\hat{y} = \sigma(z) \tag{29}$$

where $\mathrm{Pool}(\cdot)$ denotes temporal pooling over the attended feature sequence, $\mathbf{W}_o$ and b_o denote the learnable output-layer weight and bias of the TCN-SA classifier, z denotes the output logit, and $\hat{y}$ denotes the predicted probability for the positive class. The model was optimized using binary cross-entropy with logits loss as Eq. (30):

$$L = BCEWithLogitsLoss(z, y) \tag{30}$$

where y denotes the ground-truth class label. For the PCA representation, the TCN used three dilated convolutional blocks with channel widths of 32, 64, and 64, a kernel size of 7, and dilation rates of 1, 2, and 4. The self-attention block used four heads with an embedding dimension of 64. For the MS-PCA representation, a lighter configuration was adopted, with channel widths of 16, 32, and 32, a kernel size of 5, and the same dilation schedule. In this branch, the self-attention block used four heads with an embedding dimension of 32. Each convolutional block was followed by batch normalization, rectified linear unit activation, and dropout, and a 1×1 convolution was used in the residual shortcut when channel alignment was required [29], [30]. The configurations are summarized in Table 2. The output of the TCN-SA block was pooled over time and projected to a single output logit. Training was performed using the Adamax optimizer with a learning rate of 1×10^{-4}, a batch size of 64, and 50 epochs. Binary cross-entropy with logits loss was used as the training objective in both feature-representation settings.

The PCA and MS-PCA branches used branch-specific model configurations, as summarized in Table 1 and Table 2. The MS-PCA-based models used lighter configurations because the multiscale representation already combined complementary components from multiple temporal smoothing scales, and a more compact downstream model was used to reduce the risk of over-parameterization in the limited cohort. These configurations were fixed before the final subject-wise cross-validation evaluation, and the held-out test folds were not used for hyperparameter selection, early stopping, or model selection. Therefore, the comparison in this study should be interpreted as a predefined pipeline-level comparison between PCA-based and MS-PCA-based schemes rather than as a strict capacity-matched ablation isolating the independent effect of the feature representation alone. This design choice is acknowledged as a methodological limitation, and future work should include identical-capacity architectures and systematic hyperparameter optimization for all schemes.

I. Evaluation and Statistical Analysis

Model performance was evaluated using subject-wise 5-fold cross-validation, with each test fold containing data from participants not seen during training in order to reduce subject-identity leakage [9]. For each configuration, classification performance was summarized using accuracy, precision, recall, F1-score, and specificity [24]. Confusion matrices and training trajectories were also examined to describe error balance and optimization behavior. The evaluation equations are included to make explicit how window-level classification performance was computed from the aggregated confusion matrices. TP, TN, FP, and FN were obtained from the held-out subject-wise test windows, and the resulting metrics were summarized across the five folds. The window-level evaluation metrics in Eqs. (31)–(35) follow standard binary classification performance definitions [24]. Let TP, TN, FP, and FN denote the numbers of true positives, true negatives, false positives, and false negatives, respectively. The evaluation metrics were defined as

$$\mathrm{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{31}$$

$$\mathrm{Precision} = \frac{TP}{TP + FP} \tag{32}$$

$$\mathrm{Recall} = \frac{TP}{TP + FN} \tag{33}$$

$$\mathrm{F1\text{-}score} = \frac{2 \cdot \mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{34}$$

$$\mathrm{Specificity} = \frac{TN}{TN + FP} \tag{35}$$

To complement the window-level evaluation, subject-level performance was also assessed by aggregating the window-level predictions within each held-out participant. Because each test fold contained one ASD participant and one TD participant, the fold-wise window counts corresponded to the held-out subject counts within each class. For a given subject, the predicted subject label was assigned using majority voting across all windows from that subject. Specifically, if most windows from a subject were predicted as ASD, the subject-level prediction was ASD; otherwise, the subject-level prediction was TD. The subject-level vote proportion was calculated as the number of windows assigned to the majority class divided by the total number of windows from that subject. This subject-level aggregation was used only for descriptive reporting and was not used for

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

hyperparameter tuning or threshold optimization. The inferential analysis focused on two a priori planned contrasts: PCA versus MS-PCA within the RNN-BiLSTM architecture, and PCA versus MS-PCA within the TCN-SA architecture. Because the paired comparisons were obtained from the same subject-wise folds, the fold-wise test accuracies were treated as paired observations. Specifically, the statistical test used the five-fold-wise test accuracy values obtained from subject-wise 5-fold cross-validation for each scheme. Given the small number of folds and the bounded nature of accuracy values, statistical significance was assessed using the exact Wilcoxon signed-rank test rather than a parametric test. The significance level was set at $\alpha = 0.05$. The following statistical equations define the paired fold-wise comparison between PCA and MS-PCA within each temporal architecture. They were used to test whether the descriptive differences in fold-wise accuracy were supported by inferential evidence. The paired-difference definition in Eq. (36) and the rank-biserial correlation in Eq. (37) were used to support the exact Wilcoxon signed-rank comparison of fold-wise accuracies, following standard nonparametric paired-comparison and effect-size reporting procedures [31]. For a given architecture, let $A_f^{\text{MS-PCA}}$ and A_f^{PCA} denote the test accuracies in fold f, where $f = 1, \ldots, 5$. The paired difference was defined as:

$$d_f = A_f^{\text{MS-PCA}} - A_f^{\text{PCA}} \tag{36}$$

The null hypothesis stated that the median of the paired differences was equal to zero. In addition to the exact Wilcoxon signed-rank p-value, the effect magnitude was quantified using the rank-biserial correlation (r_{rb}), defined as

$$r_{rb} = \frac{W^+ - W^-}{W^+ + W^-} \tag{37}$$

where W^+ and W^- denote the sums of positive and negative signed ranks, respectively. Positive r_{rb} values indicate an advantage for the MS-PCA configuration, whereas negative values indicate an advantage for the PCA configuration.

J. Implementation and Reproducibility Details

All signal processing, feature extraction, model training, and evaluation procedures were implemented in Python 3.12.12. Numerical computation and signal processing were performed using NumPy 2.0.2, Pandas 2.3.3, and SciPy 1.16.3. Machine-learning utilities, cross-validation splitting, and evaluation metrics were implemented using scikit-learn 1.6.1, while the RNN-BiLSTM and TCN-SA models were implemented using PyTorch 2.10.0+cu128. Experiments were executed on Kaggle Notebook under Linux 6.6.122+ x86_64 with glibc 2.35, using an Intel Xeon CPU at 2.00 GHz, 4 CPU cores, 31.35 GB RAM, and two NVIDIA Tesla T4 GPUs, each with 15,360 MiB memory. PyTorch CUDA 12.8 and cuDNN 9.10.2 were used for GPU acceleration, with NVIDIA driver version 580.105.08.

To improve reproducibility, the same subject-wise fold assignments were used across all evaluated schemes. Randomness was controlled by fixing the random seed to 42 for Python, NumPy, and PyTorch. When GPU computation was used, the corresponding CUDA random seed was also fixed. Deterministic computation settings were enabled where supported, although minor numerical variation may still occur because of hardware- and backend-dependent GPU operations. The hyperparameters reported in Table 1 and Table 2 were fixed before the final subject-wise cross-validation evaluation. No held-out test fold was used for hyperparameter tuning, early stopping, or model selection. Therefore, the reported results reflect fixed-configuration evaluation under the predefined leakage-aware subject-wise cross-validation protocol rather than test-fold-driven hyperparameter optimization. Training/evaluation runtime was estimated from the execution metadata of the final Kaggle Notebook cells used for subject-wise 5-fold cross-validation. The total runtime for the 5-fold evaluation was 61.84 min for PCA + RNN-BiLSTM, 4.66 min for MS-PCA + RNN-BiLSTM, 18.63 min for PCA + TCN-SA, and 23.41 min for MS-PCA + TCN-SA, corresponding to average runtimes of 12.37, 0.93, 3.73, and 4.68 min per fold, respectively. For the MS-PCA + TCN-SA scheme, the reported runtime reflects the end-to-end execution time of the corresponding K-fold implementation cell, including preprocessing and feature-generation steps contained in the same cell.

III. Results

This section presents the experimental results of the four evaluated EEG classification schemes using held-out subject-wise 5-fold cross-validation. Training curves were used to describe optimization behavior during model fitting, whereas classification performance and statistical comparisons were assessed using held-out test folds. The results are presented in three parts: training dynamics, classification performance, and statistical comparison between PCA and MS-PCA.

A. Training Dynamics

The training dynamics of the four evaluated schemes were examined using fold-wise training accuracy and training loss trajectories over 50 epochs. Because no separate validation set was used, Fig. 9 and Fig. 10 are interpreted as optimization summaries rather than as

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

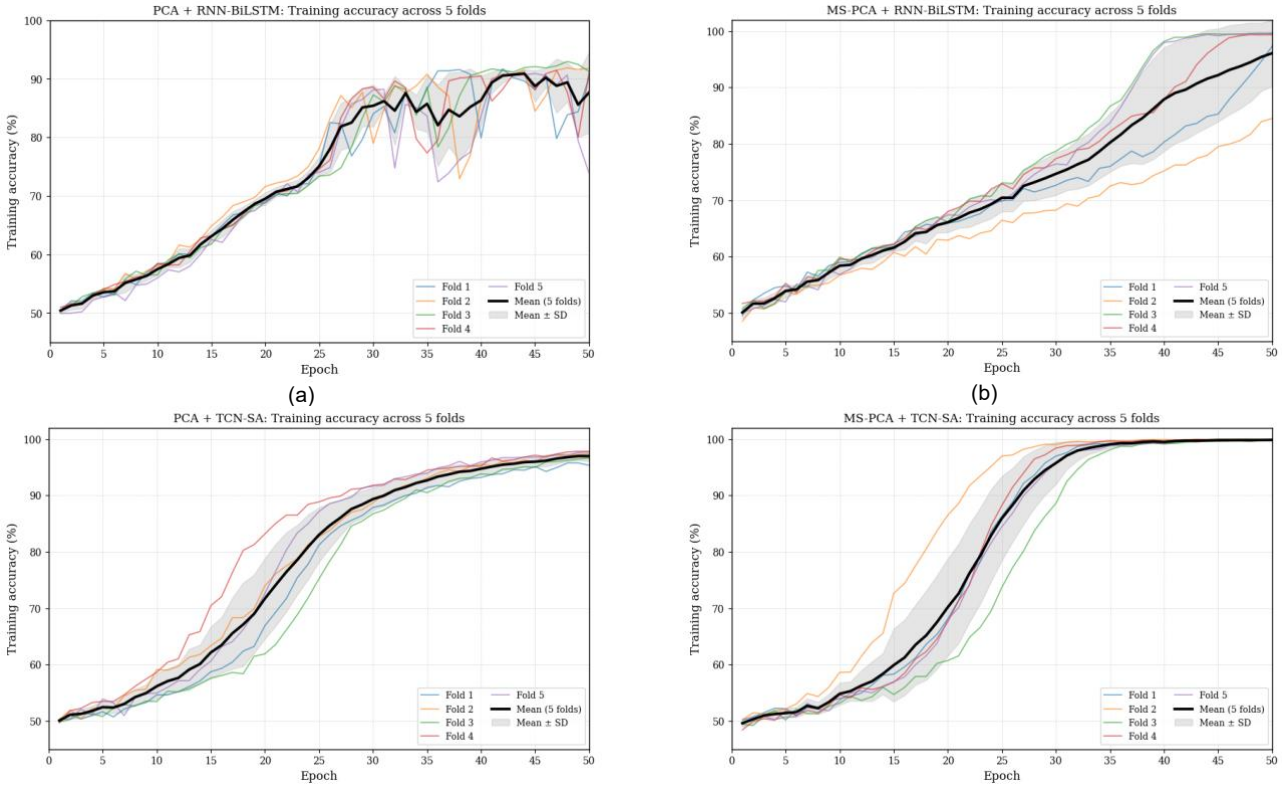

Fig. 9. Training accuracy curves for the four evaluated schemes. The figure compares fold-wise optimization behavior across (a) PCA + RNN-BiLSTM, (b) MS-PCA + RNN-BiLSTM, (c) PCA + TCN-SA, and (d) MS-PCA + TCN-SA, showing that the MS-PCA-based schemes, particularly MS-PCA + TCN-SA, achieved faster and more stable descriptive convergence.

final evidence of generalization. Final performance was assessed using held-out subject-wise test folds. As shown in Fig. 9, all schemes showed increasing training accuracy across epochs, but their convergence patterns differed. PCA + RNN-BiLSTM exhibited the slowest and least stable trajectory, whereas MS-PCA + RNN-BiLSTM showed a clearer increase in training accuracy. The TCN-SA-based schemes showed smoother fold-wise convergence than the recurrent model, with MS-PCA + TCN-SA producing the most consistent training trajectories. The training loss curves in Fig. 10 showed a similar pattern. PCA + RNN-BiLSTM produced the most fluctuating loss trajectory, whereas MS-PCA + RNN-BiLSTM showed a more stable downward trend. The TCN-SA-based schemes showed more monotonic loss reduction, particularly when combined with MS-PCA. Overall, the training curves suggest more stable optimization for the MS-PCA-based schemes, especially MS-PCA + TCN-SA, while final generalization is evaluated separately using held-out test performance.

B. Classification Performance

The aggregated confusion matrices of the four evaluated schemes are presented in Fig. 11, and the corresponding classification metrics are summarized in Table 3. Each subject-wise test fold contained one ASD participant and one TD participant, with ASD/TD window counts of 455/439, 446/448, 450/444, 440/454, and 444/450 from Fold 1 to Fold 5, respectively. Thus, the complete 5-fold evaluation included 4,470 held-out windows, with 2,235 windows for each class. Using ASD as the positive class, the aggregated confusion-matrix counts were TP = 2082, TN = 2022, FP = 213, and FN = 153 for PCA + RNN-BiLSTM; TP = 2164, TN = 2188, FP = 47, and FN = 71 for MS-PCA + RNN-BiLSTM; TP = 2171, TN = 1914, FP = 321, and FN = 64 for PCA + TCN-SA; and TP = 2183, TN = 2196, FP = 39, and FN = 52 for MS-PCA + TCN-SA. These counts indicate that the MS-PCA-based schemes descriptively reduced both false-positive and false-negative errors compared with their PCA-based counterparts. As shown in Table 3, MS-PCA + TCN-SA yielded the highest mean metric values, with an accuracy of 97.96 ± 2.37%, precision of 98.37 ± 3.40%, recall of 97.66 ± 4.39%, F1-score of 97.93 ± 2.43%, and specificity of 98.27 ± 3.63%. MS-PCA + RNN-BiLSTM showed the second-highest descriptive performance, while the PCA-based schemes showed lower mean metric values. These results indicate a

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

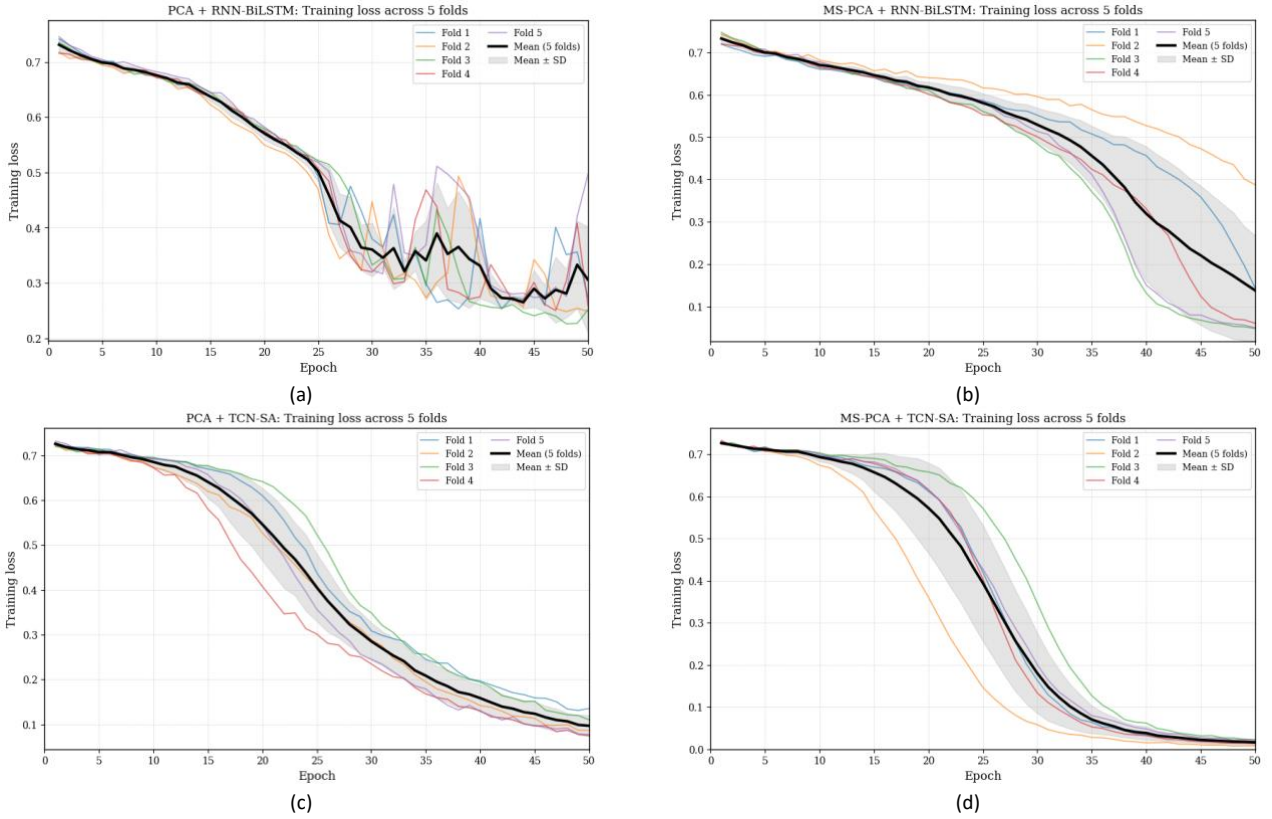

Fig. 11. Training loss curves for the four evaluated schemes. The figure compares fold-wise loss reduction during training and shows that MS-PCA + TCN-SA suggests descriptively more stable loss reduction, whereas PCA + RNN-BiLSTM showed less stable optimization behavior.

descriptive trend favoring MS-PCA-based representations; however, inferential evidence is evaluated separately in Section III.C. Subject-level aggregation was performed using majority voting over the window-level predictions of each held-out participant. All four schemes correctly classified all 10 held-out participants across the five folds. The corresponding majority-vote proportions ranged from 85.33% to 96.89% for PCA + RNN-BiLSTM, 88.79% to 99.78% for MS-PCA + RNN-BiLSTM, 70.16% to 100.00% for PCA + TCN-SA, and 89.91% to 100.00% for MS-PCA + TCN-SA. However, because this subject-level result was based on only 10 participants, it should be interpreted as a descriptive subject-level summary rather than evidence of robust clinical generalizability.

C. Statistical Comparison of PCA and MS-PCA

To determine whether the descriptive performance differences in Fig. 11 and Table 3 were supported by inferential evidence, paired nonparametric comparisons were conducted between PCA and MS-PCA within each

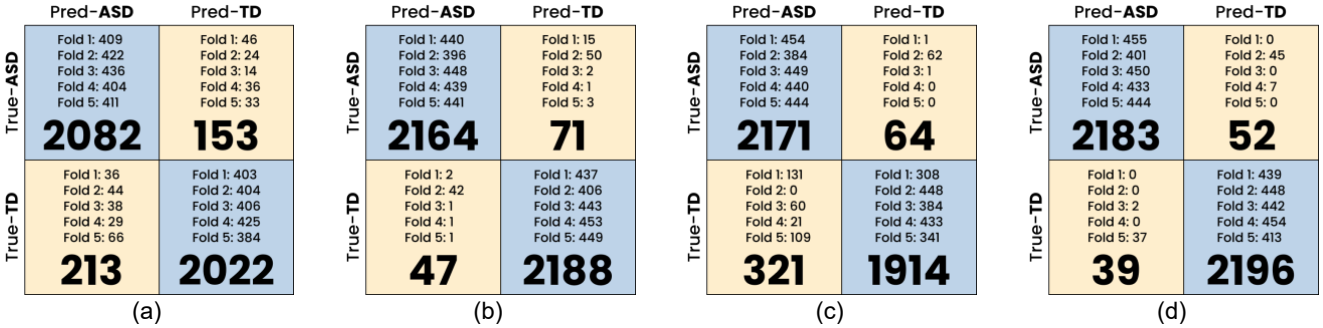

Fig. 10. Aggregated confusion matrices of the four classification schemes on the held-out test data across subject-wise 5-fold cross-validation: (a) PCA + RNN-BiLSTM, (b) MS-PCA + RNN-BiLSTM, (c) PCA + TCN-SA, and (d) MS-PCA + TCN-SA.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

Table 3. Classification performance across 5-fold cross-validation (mean ± std).

Scheme	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Specificity (%)
PCA + RNN-BiLSTM	91.81 ± 2.01	90.78 ± 2.76	93.16 ± 2.69	91.92 ± 1.94	90.47 ± 3.12
MS-PCA + RNN-BiLSTM	97.36 ± 4.33	97.86 ± 4.16	96.83 ± 4.66	97.34 ± 4.38	97.90 ± 4.07
PCA + TCN-SA	91.39 ± 4.90	88.31 ± 9.57	97.13 ± 6.17	92.04 ± 4.05	85.56 ± 12.63
MS-PCA + TCN-SA	97.96 ± 2.37	98.37 ± 3.40	97.66 ± 4.39	97.93 ± 2.43	98.27 ± 3.63

temporal architecture. As summarized in Table 4, the exact Wilcoxon signed-rank test was applied to the 5 fold-wise test accuracy values, and the rank-biserial correlation (r_{rb}) was reported as the effect-size measure. For RNN-BiLSTM, the comparison between PCA and MS-PCA yielded $W = 14$, $p = 0.1250$, and $r_{\mathrm{rb}} = 0.87$, indicating a large positive descriptive effect in favor of MS-PCA, but no statistically significant difference at the 0.05 level. For TCN-SA, the comparison yielded $W = 15$, $p = 0.0625$, and $r_{\mathrm{rb}} = 1.00$, indicating that all fold-wise differences favored MS-PCA, although the two-sided exact test remained slightly above the conventional significance threshold. Overall, the statistical comparison is consistent with a descriptive trend favoring MS-PCA within both temporal architectures, but it does not provide statistically significant evidence of MS-PCA superiority over PCA. Therefore, the observed differences should be interpreted as preliminary descriptive findings under the evaluated configurations, and statistical confirmation requires validation using a larger cohort.

IV. Discussion

The present study evaluated PCA and MS-PCA feature representations combined with RNN-BiLSTM and TCN-SA models for EEG-based ASD classification under a leakage-aware subject-wise cross-validation protocol. The results suggested a descriptive trend favoring MS-PCA-based schemes, particularly MS-PCA + TCN-SA. However, the Wilcoxon signed-rank tests did not reach statistical significance at the 0.05 level. Therefore, the observed differences should be interpreted as preliminary descriptive findings rather than statistically confirmed superiority of MS-PCA over PCA.The descriptive trend favoring MS-PCA may be related to the multiscale nature of EEG signals. Conventional PCA provides a compact single-scale representation and is useful for reducing redundancy, but it may compress localized temporal variations in nonstationary EEG data. In contrast, MS-PCA extracts components from several smoothed temporal scales and concatenates them into a compact pseudo-channel representation. This procedure may provide temporal learning models with complementary signal structures that are less accessible from a single-scale PCA representation [11], [13], [15]. The TCN-SA architecture appeared to benefit most from the MS-PCA representation. Dilated temporal convolutions allow TCNs to capture broader temporal contexts without reducing sequence resolution, while self-attention can emphasize informative temporal regions within EEG windows [23], [24]. These mechanisms may explain why the MS-PCA + TCN-SA scheme showed the most favorable descriptive pattern among the evaluated configurations. Nevertheless, this interpretation remains descriptive because the study used a limited cohort and branch-specific model configurations.

The PCA + TCN-SA result also highlights an important error-balance issue. Although this scheme achieved high recall, it produced more false-positive predictions than the MS-PCA + TCN-SA scheme. In screening-oriented contexts, a high false-positive rate may increase the burden of follow-up assessment. The lower false-positive count observed after replacing PCA with MS-PCA suggests that the multiscale representation may help the TCN-SA model better separate ASD-related patterns from control variability. However, this observation should be interpreted cautiously because it was derived from a small subject-wise cross-validation setting. The inferential analysis provides an important caution. Although the fold-wise trends favored MS-PCA in both temporal architectures, the exact Wilcoxon signed-rank tests did not reach statistical significance. The positive rank-biserial correlations suggest a consistent descriptive direction, especially for the TCN-SA comparison, but the analysis was based on only five subject-wise folds. Therefore, the findings should not be interpreted as evidence that MS-PCA is statistically superior to PCA. Larger cohorts are required to determine whether the observed descriptive trend remains reliable and generalizable.

Table 4. Exact Wilcoxon signed-rank comparison of PCA and MS-PCA based on fold-wise test accuracies.

Comparison	(n)	Statistic (W)	(p)-value	Effect size (r_{rb})	Conclusion
MS-PCA + RNN-BiLSTM vs. PCA + RNN-BiLSTM	5	14	0.1250	0.87	Not significant
MS-PCA + TCN-SA vs. PCA + TCN-SA	5	15	0.0625	1.00	Not significant

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

From a clinical perspective, the present findings should be interpreted as preliminary methodological evidence rather than as evidence of diagnostic readiness or clinical deployment. The dataset included only 10 participants, the fold-wise statistical comparisons between PCA and MS-PCA did not reach statistical significance, and detailed individualized clinical metadata were not available as complete structured variables for all participants. Therefore, the proposed framework should not be used to support clinical diagnosis, screening decisions, or patient-level decision-making in its current form. Its implication is limited to showing that leakage-aware evaluation and multiscale component-based EEG representation may be useful directions for future EEG-based ASD research. Clinical translation would require larger independent cohorts, richer clinical characterization, prospective validation, standardized artifact-control procedures, and comparison with established clinical assessment workflows.

These results are broadly consistent with previous work showing that feature construction and model

Table 5. Comparison of representative EEG-based ASD classification studies with the proposed framework.

Study	Sample Size	Features	Classifier	Validation Strategy / Leakage Control	Results	Key Comparability Notes
Rogala et al. (2023) [12]	18 ASD, 18 TD	Spectral power, PLV, and ciPLV EEG connectivity features	L2-regularized logistic regression	Nested CV with 6 outer and 5 inner splits; additional LOO subject-level prediction	Acc 0.86 using ciPLV	Small retrospective clinical cohort; high-dimensional connectivity features relative to sample size; no independent external validation
Ke et al. (2024) [7]	246 ASD, 42 controls	Spectral power, spectral coherence, connectivity, complexity, and time-domain EEG features	Ridge logistic regression with feature selection	Nested LOOCV	ROC-AUC 0.65	Class imbalance; modest discrimination; no independent external test set; possible inflated performance risk.
Tawhid et al. (2021) [32]	KAU EEG dataset, 16 subjects	STFT spectrogram EEG images	CNN-based deep learning model	70% training, 15% validation, 15% testing; subject-wise leakage control not clearly equivalent to the present protocol	Acc 99.15%	Small dataset; EEG-derived image split may risk subject-identity leakage if subject-wise separation is not explicitly enforced; no external validation.
Torghabeh et al. (2024) [33]	34 ASD, 11 normal EEG samples	Scalogram EEG images	Hybrid CNN-LSTM deep transfer learning	Older-than-5 subjects for training; younger subjects for validation	Acc 99.50% / 98.43%	Class imbalance; age-based validation; limited external generalization.
Xu et al. (2024) [25]	30 ASD, 30 TD	Time-series maps of brain functional connectivity	Combined CNN-LSTM model	Subject-wise 10-fold CV	Acc 95.60%	Complex connectivity-map pipeline; performance depends on connectivity-map construction; external validation still limited.
Present study	10 participants, 5 ASD and 5 TD	PCA and MS-PCA temporal component features	RNN-BiLSTM and TCN-SA	Leakage-aware subject-wise 5-fold CV	Acc 97.96 for MS-PCA + TCN-SA	Small cohort; non-significant Wilcoxon tests; branch-specific model configurations; no classical baseline classifiers.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

choice are important in EEG-based classification pipelines [11], [16]. They also align with studies indicating that structured EEG representations, including graph-based and connectivity-based features, can support discriminative learning in neurodevelopmental applications [17], [18], [19], [25]. The present study extends these observations by evaluating a compact multiscale PCA representation paired with two temporal modeling families under a leakage-aware subject-wise validation protocol.

Table 5 contextualizes the present framework relative to representative EEG-based ASD classification studies. The comparison emphasizes methodological differences in sample size, feature representation, classifier design, validation strategy, leakage-control consideration, reported results, and key limitations. Because the compared studies used different participant cohorts, preprocessing procedures, feature types, validation protocols, and performance-reporting levels, the reported accuracy values should not be interpreted as a direct ranking of model superiority. Although Table 5 summarizes reported performance values, direct accuracy comparison across EEG-based ASD studies is not methodologically fair. First, the dataset sizes and class distributions differ substantially across studies, ranging from small cohorts with fewer than 20 participants to larger, more imbalanced datasets. Small datasets may produce unstable fold-wise estimates, whereas imbalanced datasets may affect accuracy, precision, recall, and ROC-AUC differently. Second, validation strategies differ across studies, including nested cross-validation, leave-one-out cross-validation, fixed training-validation-testing splits, age-based validation, and subject-wise cross-validation. These protocols do not provide the same level of evidence for generalization. Third, leakage control is not equivalent across studies. In EEG classification, segment-level or image-level splitting can overestimate performance if data derived from the same participant appears in both training and testing partitions. Therefore, studies that do not explicitly perform subject-wise separation before feature fitting, windowing, normalization, and model evaluation are not directly comparable with leakage-aware subject-wise protocols. Fourth, the feature representations are fundamentally different, including spectral features, connectivity features, STFT spectrograms, scalogram images, functional-connectivity maps, and PCA/MS-PCA temporal component features. Because these representations capture different aspects of EEG activity and require different preprocessing and model assumptions, their reported accuracy values should be interpreted as context-dependent results rather than direct evidence that one method is universally superior. For these reasons, Table 5 is intended to provide a methodological comparison rather than a direct ranking of accuracy.

Several limitations should be acknowledged. The main limitation is the small sample size, consisting of only 10 participants. Although subject-wise 5-fold cross-validation was used to reduce subject-identity leakage, each test fold contained only two participants, which limits the stability of fold-level estimates and reduces statistical power. In addition, detailed individualized clinical metadata, including standardized ASD severity scores, cognitive level, medication status, and comorbidity profiles, were not available as complete structured variables for all participants. These limitations restrict clinical interpretation and highlight the need for larger datasets with richer demographic and clinical characterization. Another limitation concerns acquisition and preprocessing. The analysis was restricted to resting-state, eyes-open EEG recorded using a 16-channel acquisition setup. Therefore, the findings may not directly generalize to other acquisition protocols, channel configurations, or task conditions. Although band-pass filtering, signal-quality screening, and amplitude-based rejection were applied, the preprocessing pipeline did not include dedicated ocular or muscular artifact correction using EOG/EMG channels, ICA, or regression-based procedures. Residual artifacts may therefore have influenced the extracted EEG representations.

The study also has methodological limitations related to model comparison. The analysis focused on PCA and MS-PCA and did not include alternative feature families such as connectivity-based, graph-based, or time-frequency representations. In addition, classical baseline classifiers such as logistic regression, support vector machine, random forest, XGBoost, and conventional CNN models were not evaluated. Another important limitation is that the PCA and MS-PCA branches used branch-specific model configurations, with lighter configurations for the MS-PCA-based models. Therefore, network capacity is a potential confounding factor, and the observed differences in descriptive performance cannot be attributed solely to the MS-PCA feature representation. Instead, the results should be interpreted as the performance of the evaluated MS-PCA-based pipeline under the specified model configurations, rather than as a strictly capacity-matched ablation that isolates the independent effect of MS-PCA alone. Future studies should include classical and shallow learning baselines, identical-capacity or capacity-matched architectures, systematic hyperparameter optimization, and ablation experiments under the same leakage-aware subject-wise validation protocol.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

V. Conclusion

This study compared PCA and MS-PCA feature representations, combined with RNN-BiLSTM and TCN-SA models, for EEG-based ASD classification under a leakage-aware subject-wise 5-fold cross-validation protocol. The MS-PCA + TCN-SA scheme showed the best descriptive performance among the evaluated configurations, achieving an accuracy of 97.96 ± 2.37% with balanced precision, recall, F1-score, and specificity. However, the evidence remains preliminary because the dataset included only 10 participants, the exact Wilcoxon signed-rank comparisons between PCA and MS-PCA did not reach statistical significance, and the evaluated schemes used branch-specific model configurations without classical baseline classifiers. Therefore, the findings should not be interpreted as diagnostic readiness or statistically confirmed superiority of MS-PCA, but rather as an initial indication that multiscale component-based EEG representation may be useful for temporal deep learning. Future work should validate the framework using larger, independent, and clinically diverse datasets, with stronger artifact control, capacity-matched model comparisons, classical baseline classifiers, and external validation.

Funding

This research was funded by Universitas Syiah Kuala and Kemdiktisaintek under Grant No. 113/C3/DT.05.00/PL/2025.

Ethical Approval

This study was conducted in accordance with ethical standards for research involving human participants. The study protocol was approved by the Health Research Ethics Committee of the Faculty of Medicine, Universitas Syiah Kuala, Indonesia, with approval number 117/EA/FK/2024.

Consent to Participate

Written informed consent was obtained from all participants or from their legal guardians before data collection.

Consent for Publication

Consent for publication was obtained from all participants or their legal guardians. The manuscript does not include identifiable personal information, facial images, or individual clinical details that could reveal participant identity.

Data Availability

The datasets generated and analyzed during the current study are not publicly available due to ethical and privacy restrictions related to human EEG recordings. The data may be made available from the corresponding author upon reasonable request and subject to appropriate ethical approval and institutional permission.

Code Availability

The source code used for preprocessing, feature extraction, model training, and evaluation is available from the corresponding author upon reasonable request.

Author Contributions

Muliyadi Muliyadi contributed to study design, EEG data processing, model implementation, experimental analysis, interpretation of results, and manuscript preparation. Melinda Melinda supervised the study and contributed to conceptualization, research design, methodological validation, and manuscript revision. Yuwaldi Away contributed to research supervision, EEG acquisition protocol, methodological review, and technical validation. Syahrul Gazali contributed to the clinical aspects of the study, participant-related considerations, and interpretation of the autism spectrum disorder context. Aufa Rafiki contributed to data preparation, computational experiments, visualization, and result organization. W. K. Wong contributed to methodological review, interpretation of findings, manuscript refinement, and critical revision. All authors reviewed and approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Competing Interests

The authors declare no competing interests.

Acknowledgment

The authors acknowledge Universitas Syiah Kuala for institutional research support. The authors also thank the Department of Electrical and Computer Engineering, Faculty of Engineering, Universitas Syiah Kuala, and the Department of Neurology, Faculty of Medicine, Universitas Syiah Kuala/RSUD Dr. Zainoel Abidin, Banda Aceh, for their support during EEG data acquisition and the research process. The authors thank all participants and their legal guardians for their voluntary participation in this study.

References

[1] L. Wang, B. Wang, C. Wu, J. Wang, and M. Sun, “Autism Spectrum Disorder: Neurodevelopmental Risk Factors, Biological Mechanism, and Precision Therapy,” Feb. 01, 2023, *MDPI*. doi: 10.3390/ijms24031819.

[2] A. Reilly *et al.*, “The role of machine learning in autism spectrum disorder assessment and

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

management," *Pediatric Research 2025 98:7*, vol. 98, no. 7, pp. 2503–2517, Nov. 2025, doi: 10.1038/s41390-025-04566-0.

[3] H. A. Hatim, Z. A. A. Alyasseri, and N. Jamil, "A recent advances on autism spectrum disorders in diagnosing based on machine learning and deep learning," *Artificial Intelligence Review 2025 58:10*, vol. 58, no. 10, pp. 313-, Jul. 2025, doi: 10.1007/s10462-025-11302-x.

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

[5] 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, p. 110705, 2023, doi: 10.1016/j.pnpbp.2022.110705.

[6] 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," Dec. 01, 2023, *Springer Nature*. doi: 10.1038/s41398-023-02681-2.

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

[8] M. A. Boudewyn *et al.*, "Managing EEG studies: How to prepare and what to do once data collection has begun," *Psychophysiology*, vol. 60, no. 11, Nov. 2023, doi: 10.1111/psyp.14365.

[9] G. Brookshire *et al.*, "Data leakage in deep learning studies of translational EEG," *Front. Neurosci.*, vol. Volume 18-2024, 2024, doi: 10.3389/fnins.2024.1373515.

[10] M. Melinda *et al.*, "A comprehensive EEG dataset and performance assessment for Autism Spectrum Disorder," *Sci. Rep.*, vol. 15, no. 1, Oct. 2025, doi: 10.1038/s41598-025-18934-7.

[11] A. Anuragi, D. S. Sisodia, and R. B. Pachori, "Mitigating the curse of dimensionality using feature projection techniques on electroencephalography datasets: an empirical review," *Artif. Intell. Rev.*, vol. 57, no. 3, Mar. 2024, doi: 10.1007/s10462-024-10711-8.

[12] 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, p. 21748, 2023, doi: 10.1038/s41598-023-49048-7.

[13] G. A. Buzzell, Y. Niu, S. Aviyente, and E. Bernat, "A practical introduction to EEG Time-Frequency Principal Components Analysis (TF-PCA)," *Dev. Cogn. Neurosci.*, vol. 55, Jun. 2022, doi: 10.1016/j.dcn.2022.101114.

[14] S. Srinivasan and S. Duela Johnson, "A novel approach to schizophrenia Detection: Optimized preprocessing and deep learning analysis of multichannel EEG data," *Expert Syst. Appl.*, vol. 246, Jul. 2024, doi: 10.1016/j.eswa.2023.122937.

[15] B. Malluhi, H. Nounou, and M. Nounou, "Enhanced Multiscale Principal Component Analysis for Improved Sensor Fault Detection and Isolation," *Sensors*, vol. 22, no. 15, Aug. 2022, doi: 10.3390/s22155564.

[16] G. Eroğlu, "Electroencephalography-Based Machine Learning for Biomarker Detection in Dyslexia and Autism Spectrum Disorder: A Comparative Review of Models, Features, and Diagnostic Utility," *Diagnostics 2025, Vol. 15, Page 2388*, vol. 15, no. 18, p. 2388, Sep. 2025, doi: 10.3390/DIAGNOSTICS15182388.

[17] F. Li *et al.*, "Recognition of autism spectrum disorder in children based on electroencephalogram network topology," *Cognitive Neurodynamics 2023 18:3*, vol. 18, no. 3, pp. 1033–1045, Apr. 2023, doi: 10.1007/S11571-023-09962-Y.

[18] W. Hu, G. Jiang, J. Han, X. Li, and P. Xie, "Regional-Asymmetric Adaptive Graph Convolutional Neural Network for Diagnosis of Autism in Children with Resting-State EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 200–211, 2024, doi: 10.1109/TNSRE.2023.3347134.

[19] M. Melinda *et al.*, "Heavy–Light Soft-Vote Fusion of EEG Heatmaps for Autism Spectrum Disorder Detection," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 8, no. 1, pp. 409–429, Jan. 2026, doi: 10.35882/jeeemi.v8i1.1377.

[20] Z. Hu, L. Chen, Y. Luo, and J. Zhou, "EEG-Based Emotion Recognition Using Convolutional Recurrent Neural Network with Multi-Head Self-Attention," *Applied Sciences (Switzerland)*, vol. 12, no. 21, Nov. 2022, doi: 10.3390/app122111255.

[21] W. Zhao *et al.*, "Residual and bidirectional LSTM for epileptic seizure detection," *Front. Comput. Neurosci.*, vol. 18, 2024, doi: 10.3389/fncom.2024.1415967.

[22] S. Mohsen, S. S. M. Ghoneim, M. S. Alzaidi, A. Alzahrani, and A. M. Ali Hassan, "Classification of Electroencephalogram Signals Using LSTM and SVM Based on Fast Walsh-Hadamard Transform," *Computers, Materials and Continua*, vol. 75, no. 3, pp. 5271–5286, 2023, doi: 10.32604/cmc.2023.038758.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

[23] A. H. P. Nguyen, O. Oyefisayo, M. A. Pfeffer, and S. H. Ling, "EEG-TCNTransformer: A Temporal Convolutional Transformer for Motor Imagery Brain–Computer Interfaces," *Signals*, vol. 5, no. 3, pp. 605–632, Sep. 2024, doi: 10.3390/signals5030034.
[24] L. Huang, K. Zhou, S. Chen, Y. Chen, and J. Zhang, "Automatic detection of epilepsy from EEGs using a temporal convolutional network with a self-attention layer," *BioMedical Engineering OnLine 2024 23:1*, vol. 23, no. 1, pp. 50-, Jun. 2024, doi: 10.1186/S12938-024-01244-W.
[25] 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, p. 108196, 2024, doi: 10.1016/j.cmpb.2024.108196.
[26] O. Rainio, J. Teuho, and R. Klén, "Evaluation metrics and statistical tests for machine learning," *Scientific Reports 2024 14:1*, vol. 14, no. 1, pp. 6086-, Mar. 2024, doi: 10.1038/s41598-024-56706-x.
[27] M. Melinda, P. D. Purnamasari, F. Fahmi, E. Sinulingga, M. Muliyadi, and I. A. Yamin, "Innovative Portable Quantitative EEG Data Acquisition Based Cyton Biosensing Board using PSD Analysis," in *2024 10th International Conference on Smart Computing and Communication (ICSCC)*, 2024, pp. 631–636. doi: 10.1109/ICSCC62041.2024.10690827.
[28] S. Marukatat, "Tutorial on PCA and approximate PCA and approximate kernel PCA," *Artif. Intell. Rev.*, vol. 56, no. 6, pp. 5445–5477, Jun. 2023, doi: 10.1007/s10462-022-10297-z.
[29] J. Sheng *et al.*, "A Multi-Scale Temporal Convolutional Network with Attention Mechanism for Force Level Classification during Motor Imagery of Unilateral Upper-Limb Movements," *Entropy 2023, Vol. 25, Page 464*, vol. 25, no. 3, p. 464, Mar. 2023, doi: 10.3390/E25030464.
[30] X. Shi, B. Li, W. Wang, Y. Qin, H. Wang, and X. Wang, "EEG-VTTCNet: A loss joint training model based on the vision transformer and the temporal convolution network for EEG-based motor imagery classification," *Neuroscience*, vol. 556, pp. 42–51, Sep. 2024, doi: 10.1016/j.neuroscience.2024.07.051.
[31] F. Fiel Peres, "Effect sizes for nonparametric tests," *Biochem. Med. (Zagreb).*, vol. 36, no. 1, Feb. 2025, doi: 10.11613/BM.2026.010101.
[32] 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 June, Jun. 2021, doi: 10.1371/journal.pone.0253094.
[33] 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 2023 62:2*, vol. 62, no. 2, pp. 495–503, Nov. 2023, doi: 10.1007/s11517-023-02959-z.

Author Biography

Muliyadi was born in Punteut, Lhokseumawe, on October 28, 1976. He earned his Bachelor's degree in Applied Science from Institut Teknologi Sepuluh Nopember, Surabaya, with a major in Electrical Engineering-Information Technology. He then pursued his Postgraduate studies in Electrical Engineering at Universitas Sumatera Utara, Medan. Currently, he serves as a faculty member at Politeknik Negeri Lhokseumawe (PNL). Throughout his career, he has held various professional roles at PNL, including teaching staff, practitioner, and expert consultant for both government and non-government institutions. His research interests focus on Electrical Engineering and Information Engineering, and he actively engages in research and community service. He can be contacted at muliyadi@pnl.ac.id.

Melinda was born in Bireuen, Aceh, on June 10, 1979. She received a B.Eng degree from the Department of Electrical and Computer Engineering, Faculty of Engineering, Universitas Syiah Kuala, Banda Aceh, in 2002. She completed her master's degree in the Faculty of Electrical Department, University of Southampton, United Kingdom, with a concentration in field study of Radio Frequency Communication Systems in 2009. She completed her Doctoral degree in the Department of Electrical Engineering, Faculty of Engineering, Universitas Indonesia, in February 2018. She has been with the Department of Electrical Engineering, Faculty of Engineering, Universitas Syiah Kuala since 2002. Her research interests include multimedia signal processing and fluctuation processing. She can be contacted by email at: melinda@usk.ac.id.

Yuwaldi Away was born in South Aceh, Indonesia, in 1964. He received the degree in electrical-computer engineering from the Sepuluh Nopember Institute of Technology (ITS), Indonesia, in 1988, the M.Sc. degree from Bandung Institute of Technology (ITB), Indonesia, in 1993, and

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

the Ph.D. degree in industrial computer from the National University of Malaysia (UKM), in 2000. Since 1990, he has been a Lecturer with the Department of Electrical Engineering, Faculty of Engineering, Syiah Kuala University, Indonesia. From 1996 to 2000, he was a Research Assistant and Lecturer at the National University of Malaysia, and from 2001 to 2004. Since 2007, he has been a Professor and the Head of the Research Group for Automation and Robotics Studies at Universitas Syiah Kuala. His research interests include a combination of theory and practice, including microprocessor-based systems, simulation, automation, and optimization.

Syahrul Gazali was born in Banda Aceh in 1962. He received the M.D. degree from Andalas University, Padang, in 1988, and the Ph.D. degree from Gadjah Mada University, Yogyakarta, in 2013. He joined the Faculty of Medicine at Syiah Kuala University in Banda Aceh as a Lecturer in 1989. He completed the Neurology specialist training with the University of Indonesia, Jakarta, in 1997, and earned the Stroke Consultant credential from the Indonesian Collegium of Neurology in 2012. In 2021, he obtained Neurovascular Subspecialist Certification and was promoted to a professor of Neurology. His leadership roles include Assistant Dean, two terms as Dean, and Director of RSUD Dr. Zainoel Abidin. He has been the Chair of Indonesian Neurology Collegium since 2023, and the Director of human resources, education, and research with the Mahar Mardjono National Brain Center Hospital, Jakarta, since August 2024. His research focuses on cerebrovascular disease, with numerous publications and invited talks at national and international forums, including the World Stroke Congress and the World Congress of Neurology.

Aufa Rafiki was born on April 20, 2003, in Banda Aceh, Indonesia. He received the Bachelor of Computer Engineering degree from Universitas Syiah Kuala, Banda Aceh, Indonesia, in 2025. He received the Master of Electrical Engineering degree, with a concentration in Biomedical Engineering, from Universitas Syiah Kuala through a fast-track program. He has been involved in several research projects related to EEG signal processing, artificial intelligence, deep learning, and technology-assisted early detection of autism spectrum disorder. His research interests include biomedical signal processing, EEG analysis, artificial intelligence, machine learning, deep learning, and medical technology for autism spectrum disorder detection. He can be contacted at the email address aufa35@mhs.usk.ac.id.

W. K. Wong is a highly experienced professional engineer (P.Eng) with a strong background in the telecommunications and building services industries prior to involvement in academia. He is currently the Director of an M&E consultancy firm and serves as an Associate Professor in the Department of Electrical and Computer Engineering at Curtin University Malaysia. Dr. Wong received his PhD and Master's degrees from Universiti Malaysia Sabah in 2008 and 2016, respectively. He is a registered member of the Board of Engineers Malaysia and a member of IEEE. At Curtin Malaysia, he leads the IoT Research Group, where his research focuses on biometrics, bioinformatics, sensor technology, applied machine learning, and applied optimization. Dr. Wong has published over 100 academic articles and actively contributes to the research community as a reviewer and editor for numerous reputable journals. He can be contacted via email at: weikitt.w@ieee.org.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948