Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 22 No. 7 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v22i07.61213

PAPER

DWL-ICA: End-to-End Adaptive ICA Weighting for EEG-Based Stroke and Traumatic Brain Injury Classification

Jialu Gao(✉), Norwati Mustapha, Razali Bin Yaakob, Noridayu Binti Manshor

Universiti Putra Malaysia, Selangor, Malaysia

gs67746@student.upm.edu.my

ABSTRACT

Traditional EEG-based neurological disease classification typically relies on binary Independent Component Analysis (ICA) artifact removal, discarding all non-brain physiological signals under the assumption that they lack diagnostic value. This study proposes Dynamic Weight Learner (DWL) DWL-ICA, an end-to-end trainable framework that learns adaptive ICA component weighting for stroke and traumatic brain injury (TBI) classification. The DWL generates sample-adaptive weights for seven ICA component categories, enabling selective preservation of diagnostically informative physiological signals while suppressing true noise. Evaluated on the TUH-EEG three-class benchmark dataset—the only publicly available multi-class neurological disorder EEG dataset comprising 9,276 labeled segments—DWL-ICA achieved 0.95 macro-average ROC–AUC and 83% accuracy, representing an 11-point improvement over conventional static ICA preprocessing (0.84 baseline). Controlled ablation experiments decomposed this gain: adaptive weighting contributes 4 points (improving from 0.84 to 0.88), TMN features contribute 5 points (from 0.84 to 0.89), and joint end-to-end optimization yields an additional 2-point synergistic gain (from 0.93 to 0.95), demonstrating that the learned preprocessing strategy is specifically tailored to exploit TMN's spatial-frequency representation. The learned weights assigned moderate importance to Heart (0.60) and Eye (0.47) components without explicit regularization, demonstrating that physiological signals traditionally treated as artifacts contain meaningful diagnostic information. The framework's computational efficiency and flexible deployment options make it suitable for real-time clinical applications, including portable EEG devices in resource-limited settings.

KEYWORDS

electroencephalography (EEG), independent component analysis, deep learning, stroke classification, traumatic brain injury, adaptive preprocessing

1 INTRODUCTION

Stroke and traumatic brain injury (TBI) are leading causes of disability and mortality worldwide. Rapid and accurate diagnosis is critical for effective intervention.

Gao, J., Mustapha, N., Yaakob, R. B., Manshor, N. B. (2026). DWL-ICA: End-to-End Adaptive ICA Weighting for EEG-Based Stroke and Traumatic Brain Injury Classification. *International Journal of Online and Biomedical Engineering (iJOE)*, 22(7), pp. 150–168. https://doi.org/10.3991/ijoe.v22i07.61213

Article submitted 2026-02-21. Revision uploaded 2026-04-20. Final acceptance 2026-04-20.

© 2026 by the authors of this article. Published under CC-BY.

For ischemic stroke, thrombolytic therapy must be administered within 3–4.5 hours of symptom onset to significantly improve outcomes [1]. In TBI cases, early detection and severity assessment are crucial for preventing secondary brain damage [2]. However, conventional neuroimaging (CT/MRI) requires specialized facilities, is expensive, and is not suitable for continuous monitoring or point-of-care diagnosis.

Electroencephalography (EEG) has emerged as a promising alternative due to its portability, low cost, and suitability for real-time continuous monitoring [3], [4]. Recent portable EEG advances enable point-of-care recordings, making it particularly valuable in resource-limited settings. Quantitative EEG analyses have identified distinct electrophysiological signatures associated with stroke and TBI [5], [6], [7], establishing a foundation for automated diagnostic systems. Beyond cortical activity, clinical evidence supports the diagnostic relevance of peripheral physiological signals in these conditions. Autonomic dysregulation reflected in altered cardiac activity has been established as a prognostic indicator in acute ischemic stroke, with heart rate variability changes correlating with clinical outcomes [26]. Similarly, oculomotor dysfunction—including impaired predictive visual tracking—represents a quantifiable and noninvasively measurable neurological sequela of TBI [27]. These findings motivate the hypothesis that physiological signals routinely discarded as artifacts in EEG preprocessing may in fact encode diagnostically meaningful information specific to neurological injury.

Deep learning has become the dominant paradigm for automated EEG classification, with architectures spanning convolutional networks, recurrent models, and attention-based methods demonstrating strong performance across diverse neurological and cognitive monitoring tasks [28]. A comprehensive survey of these approaches reveals that most systems employ staged pipeline architectures in which signal preprocessing, feature extraction, and classification are optimized as sequential, independent stages [28]. The artifact removal stage within such pipelines has been extensively studied: comprehensive evaluations of EEG denoising strategies have established ICA as particularly effective for separating ocular, cardiac, and muscular contributions from neural signals [29], and automated ICA component classifiers such as ICLabel [9] have made large-scale artifact rejection feasible. EEG-based signal analysis frameworks have also demonstrated utility beyond traditional clinical diagnosis, including real-time cognitive state monitoring via portable devices [30] and EEG-assisted neurodevelopmental assessment [31], with adaptive deep learning architectures extending the scope of automated classification across multiple diagnostic domains [32].

Despite these advances, EEG-based diagnostics critically depend on preprocessing strategies. Traditional approaches apply binary Independent Component Analysis (ICA) artifact removal, discarding all non-brain signals under the assumption that they lack diagnostic value [8]. This assumption underpins widely used automated artifact rejection and classification tools [9], [10]. However, recent evidence challenges this assumption. Studies show that cardiac-related and eye-related EEG signals carry clinically meaningful information about neurological disorders, as autonomic dysregulation and ocular dysfunction are well-documented consequences of stroke and TBI [26], [27]. Recent machine learning studies have demonstrated promising results for EEG-based neurological disorder classification, with stroke detection accuracies ranging from 85% to 97% [5], [11], [12] and TBI classification accuracies exceeding 94% [13]. However, a critical challenge in advancing this field is the scarcity of publicly available benchmark datasets for multi-class classification. Most existing studies focus on binary classification tasks (e.g., stroke vs. normal or TBI vs. normal) using proprietary datasets, limiting reproducibility and cross-study comparison. The three-class labeled dataset established by Caiola et al. [18] from the Temple University Hospital (TUH) EEG Corpus represents the only publicly available benchmark for simultaneous Normal/Stroke/TBI classification, enabling evaluation of differential diagnosis capabilities—a more clinically realistic scenario where the system must distinguish between multiple neurological conditions rather than performing binary screening. Despite these advances in machine learning and the availability of benchmark data, most existing EEG classification systems rely on staged pipeline architectures. A representative baseline applies static ICA for signal preprocessing, followed by Short-Time Fourier Transform (STFT) feature extraction and convolutional neural network (CNN) classification [13].

ICA decomposes EEG signals into statistically independent components (ICs) [14], which are subsequently labeled by ICLabel into seven categories: Brain, Muscle, Eye, Heart, Line Noise, Channel Noise, and Other [9]. Under the conventional binary artifact removal paradigm, only components classified as brain with high confidence are retained, while all other components are entirely discarded [10], [15]. The resulting signals are then transformed using STFT to generate time-frequency spectrograms for CNN classification. Although this approach provides a strong baseline, achieving a reported macro-average ROC-AUC of 0.84, it imposes rigid assumptions on the diagnostic relevance of non-brain signals. The staged nature of the pipeline decouples preprocessing from classification, preventing optimization of artifact-handling strategies in a task-specific manner [16].

Formally, the classification problem addressed in this study is defined as follows: given a multichannel EEG recording acquired from $C = 19$ scalp electrodes over a fixed duration, the objective is to assign each recording to one of three diagnostic classes $y \in \{\text{Normal, Stroke, TBI}\}$. Existing staged approaches treat preprocessing and classification as independent optimization problems, such that the decision of which ICA components to retain is made without knowledge of the downstream diagnostic objective. This decoupling is the fundamental limitation targeted by the proposed framework.

To address these limitations, we propose Dynamic Weight Learner (DWL)-ICA, an end-to-end trainable framework that learns adaptive ICA component weighting for stroke and TBI classification. Rather than making binary keep-or-discard decisions, DWL-ICA learns continuous weighting factors that adaptively scale the contribution of each independent component according to its relevance to the diagnostic task. The framework employs a two-level adaptive weighting strategy: type-level weights capture the diagnostic importance of the seven ICA component categories, while component-level weights are derived by combining these type-level weights with ICLabel's probabilistic outputs for each of the 19 independent components. The DWL is implemented as a lightweight CNN that generates recording-specific type-level weights based on the unique IC composition of each EEG sample. Crucially, DWL-ICA is trained jointly with the classification network in an end-to-end manner, allowing gradients from the classification loss to directly optimize the learned weights for diagnostic performance. This joint optimization enables the discovery of preprocessing strategies tailored to the diagnostic objective, in contrast to conventional staged pipelines in which preprocessing and classification are optimized independently [16].

The primary contributions of this work are threefold. First, we present a novel end-to-end DWL-ICA architecture that integrates adaptive ICA component weighting, signal reconstruction, TMN-based feature extraction, and classification within a unified pipeline, enabling joint optimization of all components with respect to the diagnostic objective. Second, we demonstrate significant performance gains, with the complete system achieving a macro-average ROC-AUC of 0.95, representing an 11-point improvement over the 0.84 baseline. Controlled transfer learning experiments decompose this improvement: adaptive weighting alone contributes 4 points (from 0.84 to 0.88), while end-to-end optimization combined with TMN features yields an additional 7 points (from 0.88 to 0.95). Third, we show that the learned DWL weights can be extracted and applied as a plug-and-play module within existing STFT-based pipelines, yielding a 4-point performance improvement with minimal changes to the underlying infrastructure. This confirms the standalone effectiveness of adaptive weighting and supports its practical deployment in legacy systems.

The remainder of this paper is organized as follows. Section 2 presents the detailed methodology of the DWL-ICA system. Section 3 presents the experimental setup, datasets, and comprehensive results. Section 4 discusses the implications of our findings. Section 5 concludes the paper.

2 MATERIALS AND METHODS

2.1 DWL-ICA architecture overview

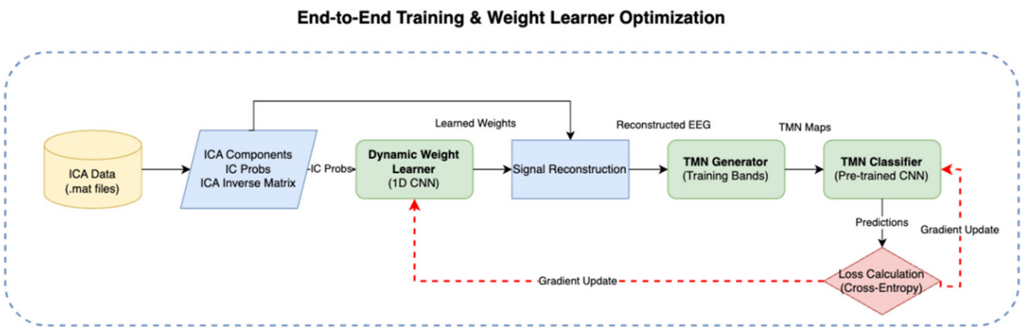

Fig. 1. DWL-ICA overall pipeline

The DWL-ICA system integrates four core components into a unified end-to-end trainable framework: (1) IC probability input from ICLabel, (2) a DWL that generates adaptive component weights, (3) weighted signal reconstruction with TMN feature extraction, and (4) a CNN-based classifier. Unlike traditional staged pipelines where preprocessing and classification are optimized independently, our architecture enables joint optimization of all components with respect to the diagnostic objective. The complete pipeline architecture is illustrated in Figure 1.

IC probability input. The input to the DWL-ICA system consists of probabilistic component classifications produced by ICLabel for each independent component. For a batch of B samples, each containing 19 independent components, ICLabel outputs a probability tensor $P \in \mathbb{R}^{(B \times 19 \times 7)}$, where each component is assigned probability scores across seven physiological categories: Brain, Muscle, Eye, Heart, Line Noise, Channel Noise, and Other.

Instead of applying hard thresholds to make binary keep-or-remove decisions, the proposed framework preserves the full probability distributions, thereby retaining uncertainty information that is critical for adaptive weighting. For instance, a component with probabilities of 0.6 for Brain, 0.3 for Muscle, and 0.1 for Eye reflects a mixture of physiological sources. Under conventional binary artifact removal, such a component would be forced into a single class, potentially discarding diagnostically informative signals.

By maintaining complete probability profiles, the DWL can learn nuanced weighting strategies that explicitly account for mixed physiological contributions. This probabilistic representation enables a transition from rigid, rule-based preprocessing to a flexible, data-driven approach. The seven ICLabel categories are summarized in Table 1.

Table 1. IC label component categories

Category	Physiological Source	Example Characteristics
Brain	Cortical neural activity	Alpha (8–13 Hz) and Beta (13–30 Hz) rhythms
Muscle	EMG contamination	High-frequency (>30 Hz) bursts
Eye	Ocular movements	Low-frequency (<4 Hz) frontal activity
Heart	Cardiac activity (ECG)	Rhythmic ~1 Hz pulses
Line Noise	Electrical interference	Sharp 50/60 Hz peaks
Channel Noise	Electrode artifacts	Spatially localized noise
Other	Unclassified sources	Mixed characteristics

2.2 DWL

The DWL is a compact 1D CNN designed to transform IC probability distributions into adaptive component weights. The network employs multi-scale convolutions to capture both local patterns within individual components and global compositional structure across the complete set of 19 ICs.

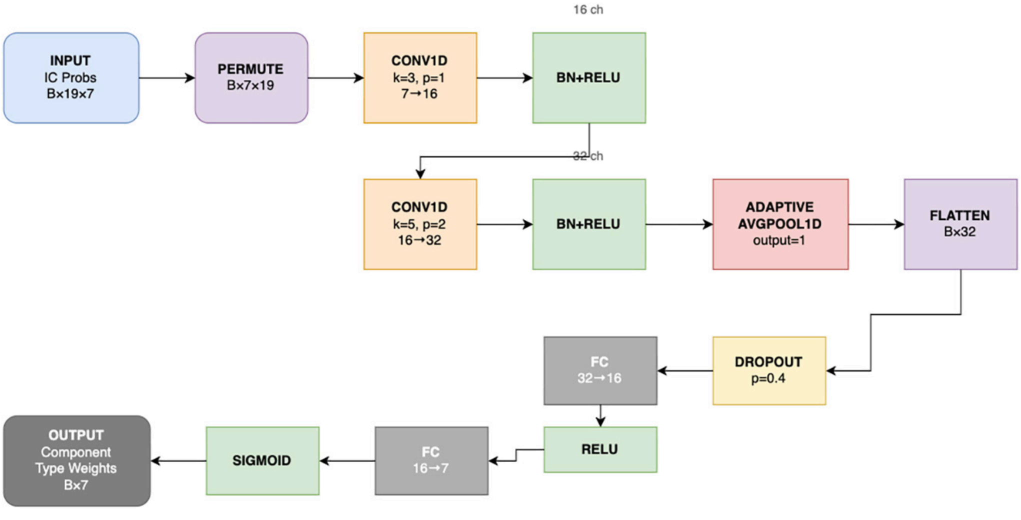

Fig. 2. DWL architecture

The DWL architecture consists of two convolutional blocks with increasing channel capacity (Conv1D: 7 to 16 with kernel size 3, followed by Conv1D: 16 to 32 with kernel size 5), each followed by batch normalization [19] and ReLU [20] activation. An adaptive average pooling layer reduces the temporal dimension to produce a 32-dimensional feature vector. This is followed by dropout [21] (p = 0.4) for regularization, then two fully connected layers (FC: 32 to 16 and 16 to 7) with ReLU activation. The final layer applies sigmoid activation to produce type-level weights W_type ∊ ℝ^(B × 7) bounded in [0,1], representing the learned diagnostic importance of each IC category. The detailed architecture is shown in Figure 2.

The DWL employs a two-level weighting mechanism. First, the network generates type-level weights W_type ∊ ℝ^(B × 7) that represent the diagnostic importance of each independent component (IC) category. These weights are sample-adaptive, such that each EEG recording is assigned a distinct set of weights based on its specific IC composition. Second, the type-level weights are combined with the IC probability distributions to derive component-level weights as follows:

$$W_{component}[i] = \sum_{j=1}^{7} P[i,j] \times W_{type}[j] \tag{1}$$

where i indexes the 19 ICs and j indexes the seven IC categories. This operation is efficiently implemented as W_component = einsum('bcl,bl->bc', P, W_type).

This two-level weighting strategy reduces the parameter space from 19 independent component weights to only 7 type-level weights, thereby improving generalization while enabling interpretable analysis of the diagnostic contribution of each IC category. In contrast to conventional binary ICA preprocessing, which applies fixed weights (Brain = 1.0, all others = 0.0), the DWL learns sample-adaptive weights through end-to-end optimization for each EEG recording.

Weighted Reconstruction and TMN Feature Extraction

After weight generation, the system performs adaptive signal reconstruction by applying the learned weights to the ICs before projecting them back into sensor space. Given ICs ∊ ℝ^(B × 19 × L) and component weights W_component ∊ ℝ^(B × 19), the weighted components are computed as:

$$\text{IC\_weighted} = \text{ICs} \odot \text{W\_component} \tag{2}$$

These weighted components are then projected back to the 19-channel sensor space using the ICA inverse matrix:

$$\text{EEG\_reconstructed} = \text{A\^{}(-1)} \times \text{IC\_weighted}. \tag{3}$$

The reconstructed EEG signals are subsequently converted into spatial-frequency representations using the Topographic Map Network (TMN) framework. Seven frequency bands are considered: delta (1–4 Hz), theta (4–8 Hz), alpha low (8–10 Hz), alpha high (10–13 Hz), beta low (13–20 Hz), beta high (20–30 Hz), and gamma (30–50 Hz). For each band, the power spectral density (PSD) is estimated via Welch's method (segment length = 2048), and the mean PSD values are extracted for each of the 19 channels. These values are then interpolated using Radial Basis Function (RBF) interpolation with cubic basis functions to generate 134×134-pixel topographic maps.

Finally, the seven topographic maps are stacked to form the complete TMN representation with dimensions B × 7 × 134 × 134, effectively capturing both the spatial distribution and spectral characteristics of neural activity.

2.3 End-to-end training

The TMN feature maps serve as input to a pre-trained CNN classifier comprising three convolutional blocks with channel dimensions of 7, 64, 64, and 128, respectively. Each block incorporates batch normalization, ReLU activation, and max pooling layers. This stage is succeeded by adaptive average pooling and a three-layer fully connected classifier consisting of 8,192, 256, 100, and 3 neurons, structured with dropout regularization. The classifier is initialized using a model pre-trained on the static ICA baseline pipeline. The entire architecture undergoes end-to-end training utilizing cross-entropy loss, with L2 regularization applied exclusively to the Brain-type weights:

$$L_{total} = L_{CE}(\hat{y}, y) + \lambda \left\| W_{type}[Brain] \right\|_2^2 \tag{4}$$

where $\lambda = 0.01$. This constraint guides Brain components toward maximum importance (weight ≈ 1.0), while weights for Eye, Heart, and artifact categories are learned purely from data without explicit regularization. This asymmetric strategy balances domain knowledge, which is brain activity that should receive maximum diagnostic importance, with data-driven discovery of physiological signal utility.

A dual learning rate strategy is employed: the pre-trained classifier is fine-tuned with learning rate 1 × 10^(–5), while the DWL trains from random initialization with learning rate 5 × 10^(–4). Both use Adam optimizer [22] ($\beta_1 = 0.9$, $\beta_2 = 0.999$). All models were trained for 80 epochs with early stopping based on validation loss. All experiments were implemented using the PyTorch framework [23].

During backpropagation, gradients from the classification loss flow backward through all components, which are the classifier, TMN generation, weighted reconstruction, and the Dynamic Weight Learner. This enables joint optimization of preprocessing and classification, allowing the DWL to learn which IC components maximize diagnostic accuracy. This contrasts with traditional staged pipelines where preprocessing decisions are made independently of classification performance.

Transfer Learning Experiment Design

To isolate the standalone contribution of the DWL, a controlled transfer learning experiment was conducted. This experiment aims to quantify the performance gains attributable specifically to the adaptive component-weighting mechanism.

In this setup, the trained DWL module is extracted from the full DWL-ICA pipeline and integrated into an STFT-based EEG classification framework. The STFT system follows the architecture proposed by Caiola et al. [18], where EEG signals are processed using the STFT and subsequently classified using a CNN, as illustrated in Figure 3 (right panel). For each 19-channel EEG recording, a one-sided STFT is computed independently per channel. The resulting real and imaginary components are concatenated to form a 129 × 256 × 38 tensor (frequency bins × time frames × channels). This spectral representation is passed through convolutional layers for feature extraction, followed by global average pooling and fully connected layers for classification.

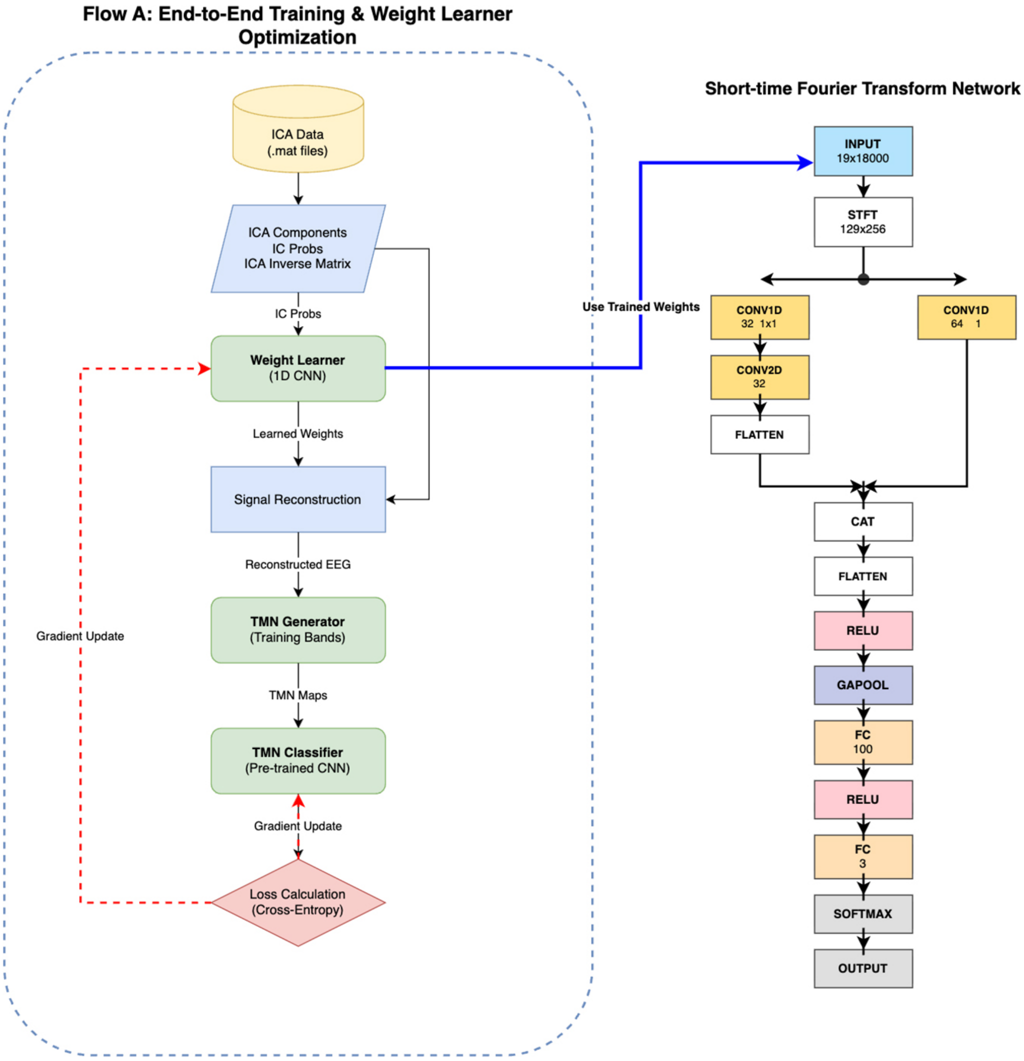

Fig. 3. Transfer learning architecture comparison

The experimental design employs matched architectures with controlled variation. Two models are trained using identical STFT feature extraction and CNN classifier architectures, differing only in their preprocessing strategy:

- A static ICA baseline with binary artifact removal
- A DWL-enhanced variant with adaptive component weighting

The DWL parameters are transferred from the pre-trained end-to-end DWL-ICA system and kept frozen during training, while the STFT-CNN components are trained normally. Importantly, the CNN classifier is retrained for both configurations to ensure a fair comparison, allowing the network to learn features optimally for each preprocessing method.

By comparing these two models, which share identical feature extraction and classification architectures and differ only in preprocessing, the experiment isolates the contribution of the DWL. The resulting difference in ROC-AUC directly quantifies the diagnostic value gained from adaptive component weighting compared to conventional binary artifact removal, which will be further explained in the next section.

3 RESULTS

3.1 Datasets

This study utilized the three-class labeled dataset established by Caiola et al. [18] from the TUH EEG Corpus [17]. This dataset represents the only publicly available benchmark for simultaneous Normal/Stroke/TBI classification, enabling reproducible evaluation and cross-study comparison. The dataset comprises clinically acquired EEG recordings from patients diagnosed with stroke or TBI, along with normal controls without neurological disorders. The dataset comprises clinically acquired EEG recordings from patients diagnosed with stroke or TBI, along with normal controls without neurological disorders. All recordings were obtained using a standardized 19-channel montage following the international 10-20 electrode placement system (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Fz, Cz, Pz), sampled at 250 Hz.

All data were collected by TUH under appropriate institutional review board approval and informed consent procedures in accordance with the Declaration of Helsinki. As this study involved only secondary analysis of fully de-identified public data, no additional ethics approval was required.

The complete dataset composition is presented in Table 2. The corpus contains recordings from 1390 subjects (498 Normal, 471 TBI, 421 Stroke) across 1571 clinical sessions. Following the preprocessing protocol detailed below, these recordings yielded 9276 non-overlapping 3-minute EEG segments that constitute the final analysis dataset. The data were partitioned using a stratified 90/10 train-test split, maintaining class distribution across both sets. This division resulted in 8349 training segments and 927 testing segments. The test set was held completely fixed throughout all experiments to ensure reproducibility and enable direct performance comparison with baseline methods.

Table 2. Dataset composition and train-test distribution

Category	Subjects	Recordings	Segments	Train	Test
Normal	498	527	3091	2782	309
TBI	471	572	3351	3016	335
Stroke	421	472	2834	2551	283
Total	1390	1571	9276	8349	927

EEG recordings underwent a standardized preprocessing pipeline to prepare data for analysis. Multi-file recording sessions were concatenated into continuous signals, with the initial 60 seconds discarded to eliminate recording stabilization artifacts. Continuous recordings were then segmented into non-overlapping 3-minute segments, with any incomplete final segments removed. All signals were bandpass filtered between 1–100 Hz using finite impulse response (FIR) filters to remove slow drifts and high-frequency noise, then re-referenced to the common average reference (CAR) across all 19 channels to minimize background electrical interference.

Independent Component Analysis was applied to each 3-minute segment using the extended Infomax algorithm [14] to decompose the 19-channel mixed signals into 19 statistically ICs. Each IC was then automatically classified using ICLabel [9], which assigns probability distributions across seven physiological categories: Brain, Muscle, Eye, Heart, Line Noise, Channel Noise, and Other.

The key methodological distinction from conventional approaches lies in how these ICA components are utilized. Traditional pipelines implement binary artifact removal, retaining only components classified as "Brain" with high confidence while completely discarding all other components [10], [15]. In contrast, our DWL-ICA framework preserves the complete ICLabel probability distributions for all 19 components without immediate removal. These probability profiles serve as input to the Dynamic Weight Learner, which learns to adaptively weight each component's contribution based on its diagnostic utility for stroke and TBI classification. This adaptive weighting strategy enables the system to discover which physiological signals contain task-relevant information rather than assuming all non-brain components are pure noise.

3.2 Overall performance comparison

Classification performance was evaluated using macro-average Receiver Operating Characteristic Area Under the Curve (ROC-AUC) [24] together with accuracy, precision, recall, and F1-score [25], following standard practice for imbalanced multi-class medical classification. The proposed DWL-ICA system achieved 0.95 macro-average ROC-AUC, representing an 11-point absolute improvement over the baseline STFT method, which reported 0.84 using static ICA preprocessing with brain-only component retention. The system additionally achieved 83% classification accuracy with balanced performance across all three diagnostic categories (refer to Table 3).

To decompose the sources of this improvement, we conducted transfer learning experiments where the trained DWL was applied to the baseline STFT feature extraction pipeline. This DWL-STFT configuration achieved 0.88 macro-average ROC-AUC and 71% accuracy, isolating the contribution of adaptive component weighting from other architectural innovations.

Table 3. Overall performance comparison

Method	Preprocessing	Features	Classifier	Macro-Avg ROC-AUC	Accuracy	AUC Improvement
Baseline [13]	Static ICA (brain-only)	STFT	CNN	0.84	–	–
DWL-STFT (Transfer)	DWL weighting	STFT	CNN	0.88	71%	+0.04
Static ICA + TMN	Static ICA (brain-only)	TMN	CNN	0.89	72%	+0.05
DWL-ICA (End-to-End)	DWL weighting	TMN	CNN	0.95	83%	+0.11

The 11-point total improvement decomposes into two distinct contributions. Adaptive weighting alone accounts for 4 points (0.84 to 0.88), validated through the transfer learning experiment where only the preprocessing strategy changed while feature extraction remained constant. The remaining 7 points (0.88 to 0.95) arise from synergistic effects of end-to-end optimization combined with TMN's spatial-frequency representation, demonstrating that joint training of preprocessing and classification components yields substantial additional gains beyond modular improvements.

3.3 Per-class performance analysis

The improvement is distributed across all three diagnostic categories, with particularly strong gains in stroke classification as shown in Table 4. Normal class ROC-AUC increased from 0.87 (baseline) to 0.95 (DWL-ICA), stroke from 0.83 to 0.97, and TBI from 0.82 to 0.94. The largest improvement of 14 points in stroke detection suggests that the adaptive weighting strategy may be particularly effective at preserving physiological signals relevant to cerebrovascular pathology.

Table 4. Per-class ROC-AUC performance comparison

Class	Baseline [13]	DWL-ICA	Absolute Improvement
Normal	0.87	0.95	+0.08
Stroke	0.83	0.97	+0.14
TBI	0.82	0.94	+0.12
Macro-Avg	0.84	0.95	+0.11

Table 5 presents comprehensive classification metrics for both the transfer learning configuration and the end-to-end system. The DWL-STFT transfer learning model achieved 71% accuracy with balanced precision and 72% recall across classes. The complete DWL-ICA system substantially improved all metrics, reaching 83% accuracy with 84% precision and 84% recall. Notably, the end-to-end system achieved relatively balanced performance across the three classes despite dataset imbalance, with per-class F1-scores ranging from 81% (TBI) to 88% (Stroke).

Table 5. Comprehensive classification metrics

Method	Class	Precision	Recall	F1-Score	Support
DWL-STFT (Transfer)	Normal	68%	80%	73%	283
	Stroke	77%	73%	75%	324
	TBI	69%	61%	65%	320
	Macro Avg	71%	72%	71%	927
	Accuracy			71%	
DWL-ICA (End-to-End)	Normal	76%	90%	83%	309
	Stroke	90%	86%	88%	283
	TBI	87%	76%	81%	335
	Macro Avg	84%	84%	84%	927
	Accuracy			83%	

3.4 Confusion matrices

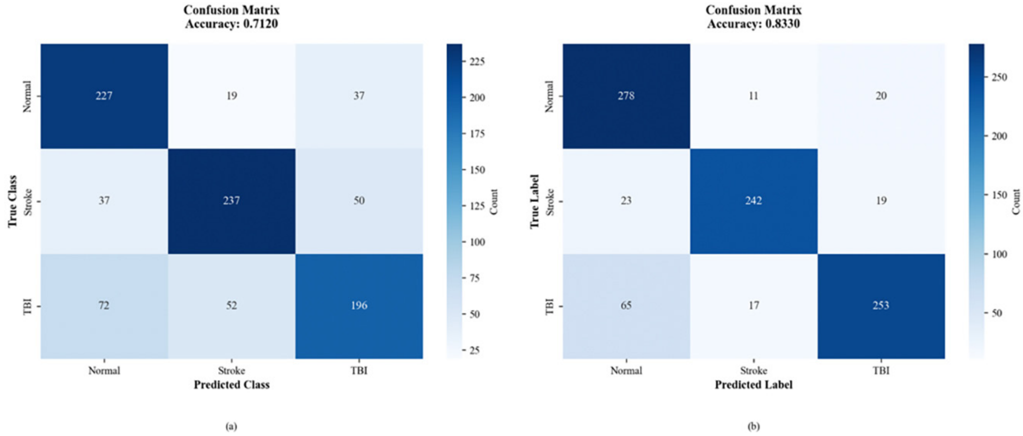

Fig. 4. Confusion matrices. (a) DWL-STFT (Transfer) (b) DWL-ICA (End-to-End)

Figure 4 presents confusion matrices for both models, revealing the nature of classification errors. The DWL-STFT transfer learning model (see Figure 4a) shows moderate confusion between all three classes, with 227 of 283 normal cases correctly classified (80% recall), but substantial misclassification of TBI cases as normal (72 cases). The end-to-end DWL-ICA system (see Figure 4b) substantially reduces these errors, correctly classifying 773 of 927 test samples (83% accuracy). Normal class recall improves to 90% (278/309), and TBI misclassification as normal drops to 65 cases. The remaining confusion occurs primarily at class boundaries, with stroke cases occasionally misclassified as TBI (19 cases) and vice versa (17 cases), likely reflecting overlapping neurophysiological signatures of these two injury types.

A systematic analysis of the 154 residual errors in the DWL-ICA system reveals two distinct failure modes. The dominant error pattern is misclassification of TBI recordings as Normal, accounting for 65 cases (42% of all errors). This is consistent with the neurophysiological profile of TBI: EEG abnormalities in TBI often manifest as non-specific diffuse slowing rather than the focal lateralized changes characteristic of stroke, making borderline TBI recordings difficult to distinguish from normal background activity. The secondary failure mode is bidirectional Stroke–TBI boundary confusion (19 Stroke misclassified as TBI and 17 TBIs misclassified as Stroke misclassifications), which is clinically interpretable: both conditions can produce overlapping patterns of diffuse cortical dysfunction, including increased delta power and reduced alpha reactivity, particularly in the subacute phase. The near-identical learned DWL weights across all three diagnostic categories (see Figure 5b) suggest that the framework learns a universal physiological signal weighting strategy rather than class-specific preprocessing, which may limit discriminability at the Stroke–TBI boundary. Future work incorporating class-conditioned weighting or adversarial training on hard boundary cases may address this limitation.

3.5 Learned weight analysis

Table 6 presents the learned type-level weights averaged across all 927 test set samples. The weights reveal a clear hierarchical pattern reflecting diagnostic importance. Brain components receive near-maximum weight (1.0000 ± 0.0001), confirming their essential role in neurological diagnosis. Physiological components exhibit moderate-to-substantial preservation: Heart components receive weight 0.5960 ± 0.0292, while Eye components receive weight 0.4748 ± 0.0457. In sharp contrast, artifact categories (Muscle, Line Noise, Channel Noise, Other) are consistently suppressed to near-zero values (0.0001 ± 0.0002), demonstrating the system's ability to discriminate between informative physiological signals and genuine noise.

Table 6. Type-level weight statistics (Test Set, n = 927)

IC Type	Mean Weight	Std Dev	Min	Max
Brain	1.0000	0.0001	0.9959	1.0000
Heart	0.5960	0.0292	0.4950	0.7077
Eye	0.4748	0.0457	0.3255	0.6521
Muscle	0.0001	0.0002	0.0000	0.0053
Line Noise	0.0001	0.0002	0.0000	0.0047
Channel Noise	0.0001	0.0002	0.0000	0.0059
Other	0.0001	0.0002	0.0000	0.0049

The remarkably low standard deviations for Brain (0.0001), Heart (0.0292), and Eye (0.0457) components indicate consistent weighting strategies across individual samples. The maximum observed weights remain bounded—Eye components never exceed 0.65 and Heart components never exceed 0.71—suggesting the learned strategy maintains Brain activity as the primary diagnostic signal while selectively incorporating supportive physiological information.

Figure 5 visualizes the weight distributions and their consistency across disease categories. Panel (a) displays mean weights with error bars for each IC type, clearly illustrating the three-tier hierarchy: Brain at maximum (1.0), physiological signals (Heart ≈ 0.60, Eye ≈ 0.47) at moderate levels, and artifacts near zero. Panel (b) compares weights across Normal, Stroke, and TBI samples, revealing nearly identical patterns for all three conditions, with differences between categories smaller than within-category variance. This cross-category consistency suggests that Heart and Eye components provide universal pathology indicators rather than disease-specific discriminative features.

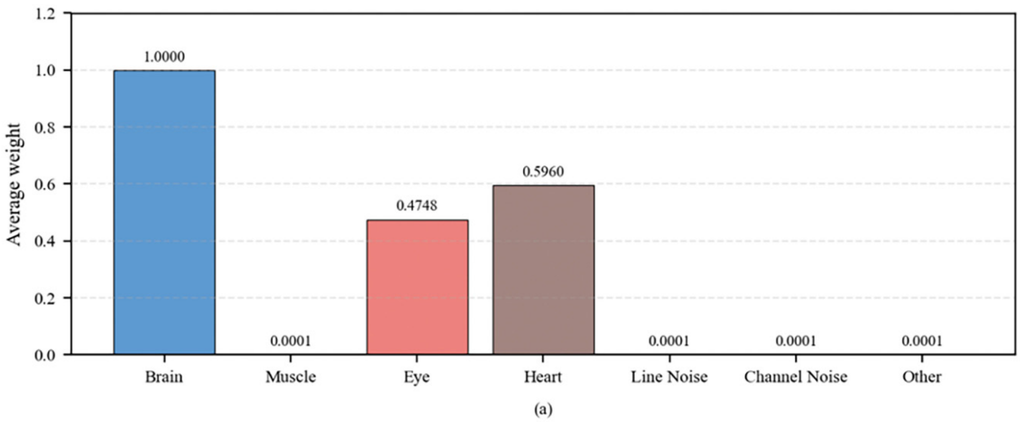

Fig. 5. (Continued)

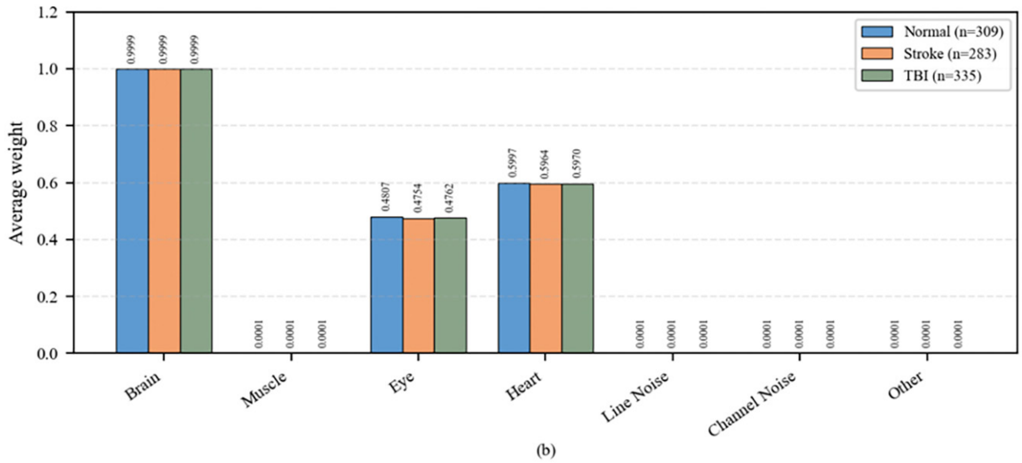

Fig. 5. Weight distribution analysis: (a) overall average weights for 7 IC classifications; (b) average weights of disease category for 7 IC classifications

4 DISCUSSION

The controlled ablation experiments provide systematic validation of the adaptive weighting hypothesis. By keeping the feature extraction method (STFT) fixed while applying adaptive preprocessing, the observed 4-point improvement in ROC-AUC (from 0.84 to 0.88) can be directly attributed to the change in preprocessing strategy, namely, the use of DWL-based adaptive weighting instead of static brain-only ICA. This isolated effect demonstrates that Eye (weight approximately 0.47) and Heart (weight approximately 0.60) components contain diagnostically relevant information that is systematically discarded under conventional binary artifact removal.

These empirical findings are supported by established clinical evidence. Ocular movement abnormalities, including saccadic dysfunction and impaired gaze control, are well-documented sequelae of both ischemic stroke and traumatic brain injury. Similarly, autonomic dysregulation reflected in altered cardiac activity has been widely observed in stroke and TBI patients, with changes in heart rate variability serving as clinically meaningful prognostic indicators. The traditional binary removal paradigm eliminates these physiological signals entirely, leading to measurable information loss, as quantified by the observed 4-point performance gap.

The relative magnitudes of the learned weights further align with clinical expectations. While Brain components receive maximal importance (weight equals 1.00), Eye and Heart components are assigned moderate weights (approximately 0.47 and approximately 0.60, respectively), reflecting their role as supplementary rather than primary diagnostic cues. Importantly, these weights emerged without explicit regularization constraints on non-brain components. Only Brain-type weights were guided toward unity via L2 regularization (λ equals 0.01), whereas Eye, Heart, and artifact categories were learned entirely from data. This data-driven emergence reinforces the interpretation that these physiological components carry genuine diagnostic information rather than reflecting imposed priors.

Moreover, the consistency of cross-category weighting patterns across Normal, Stroke, and TBI samples (Figure 5b) suggests that these physiological signals act as general indicators of neurological pathology rather than class-specific biomarkers. This observation implies that the learned weighting strategy may generalize to other neurological conditions beyond the three classes examined, although this hypothesis warrants explicit validation in future studies.

Beyond adaptive weighting alone, the additional 2-point synergistic gain in ROC-AUC achieved through end-to-end optimization (from 0.93 to 0.95) highlights the benefits of joint training. This gain emerges from the interaction between adaptive weighting (4 points) and TMN features (5 points), as their combined effect (11 points total) exceeds the sum of their individual contributions (9 points). End-to-end learning enables classification loss gradients to propagate backward through signal reconstruction, allowing the DWL to learn task-specific weights optimized directly for diagnostic performance. This co-adaptation between preprocessing and feature extraction, unattainable in conventional staged pipelines, produces gains that exceed those achievable through modular improvements alone.

Collectively, these results challenge the prevailing assumption that all non-brain physiological signals in EEG recordings should be treated as artifacts and removed entirely. Instead, they demonstrate that the distinction between "signal" and "noise" is inherently context- and task-dependent. Physiological components traditionally classified as artifacts may encode valuable diagnostic information when interpreted through an appropriate clinical and computational framework. This insight opens avenues for adaptive preprocessing strategies across a broad range of biomedical signal processing applications, such as cardiac monitoring and sleep stage classification, whereby rigid artifact removal rules may similarly discard task-relevant information.

Several limitations of the present study warrant acknowledgement. First, all experiments were conducted on a single-center dataset (TUH EEG Corpus), and the generalizability of the learned DWL weights to recordings from different clinical sites, EEG hardware configurations, or patient populations remains to be established through multi-center validation. Second, ICA decomposition quality is sensitive to segment length and signal stationarity; the 3-minute segments used here provide sufficient data for stable decomposition, but performance in shorter clinical recordings or real-time streaming contexts may differ. Third, the framework relies on ICLabel probability estimates as input to the Dynamic Weight Learner: systematic misclassification of ICA components by ICLabel—for example, conflating cardiac and brain components in recordings with pronounced pulse artifact—could propagate errors into the learned weights. Finally, the study is retrospective in design, and the clinical utility of the framework for prospective patient triage or treatment guidance has not yet been evaluated. These limitations define the scope for future work, including multi-center trials, evaluation on shorter EEG segments, and prospective clinical validation.

5 CONCLUSION

This study presented DWL-ICA, an end-to-end trainable framework for EEG-based neurological disease classification that fundamentally challenges the conventional practice of binary artifact removal in ICA preprocessing. By introducing adaptive component weighting and joint optimization with downstream classification, the proposed system achieved a macro-average ROC-AUC of 0.95 in distinguishing Normal, Stroke, and TBI cases, representing an 11-point absolute improvement over traditional static ICA-based pipelines.

Controlled ablation experiments enabled a clear decomposition of this performance gain into three sources: adaptive weighting contributes 4 points, TMN features contribute 5 points, and their joint end-to-end optimization yields an additional 2-point synergistic gain. Notably, the DWL consistently assigned substantial weights to components typically labeled as artifacts, particularly Heart (approximately 0.60) and Eye (approximately 0.47), despite the absence of explicit regularization constraints. This data-driven outcome provides physiological validation for retaining such components, aligning with established clinical evidence of ocular dysfunction and autonomic dysregulation in stroke and TBI patients.

The framework supports flexible deployment strategies: the DWL-STFT configuration can be deployed as a plug-and-play module in existing STFT-based pipelines with minimal infrastructure changes, providing a 4-point improvement, while the full DWL-ICA system achieves 0.95 AUC for scenarios requiring maximum performance. Future directions of this study include multi-center validation, expanded diagnostic coverage, and prospective clinical trials to assess impact on patient outcomes. Overall, the DWL-ICA framework demonstrates that principled integration of domain knowledge with data-driven learning can yield systems that are both high-performing and interpretable, offering a compelling pathway toward trustworthy and clinically deployable AI in healthcare.

6 ACKNOWLEDGMENT

The authors thank the Neural Engineering Data Consortium at Temple University for making the TUH EEG Corpus publicly available.

7 REFERENCES

[1] S. J. Mendelson and S. Prabhakaran, "Diagnosis and management of transient ischemic attack and acute ischemic stroke: A review," *JAMA*, vol. 325, no. 11, pp. 1088–1098, 2021. https://doi.org/10.1001/jama.2020.26867

[2] M. Vishwanath *et al.*, "Investigation of machine learning approaches for traumatic brain injury classification via EEG assessment in mice," *Sensors*, vol. 20, no. 7, 2020. https://doi.org/10.3390/s20072027

[3] F. Erani *et al.*, "Electroencephalography might improve diagnosis of acute stroke and large vessel occlusion," *Stroke*, vol. 51, no. 11, pp. 3361–3365, 2020. https://doi.org/10.1161/STROKEAHA.120.030150

[4] N. Sanossian and E. Fink, "What will the mobile stroke unit of the future look like, and will EEG have a role?" *Neurology*, vol. 101, no. 24, pp. 1085–1086, 2023. https://doi.org/10.1212/WNL.0000000000208047

[5] W. Tong, J. Zhang, F. Chen, W. Shi, L. Zhang, and J. Wan, "A novel stroke classification model based on EEG feature fusion," *Sci. Rep.*, vol. 15, 2025. https://doi.org/10.1038/s41598-025-92807-x

[6] C. Bentes *et al.*, "Quantitative EEG and functional outcome following acute ischemic stroke," *Clin. Neurophysiol.*, vol. 129, no. 8, pp. 1680–1687, 2018. https://doi.org/10.1016/j.clinph.2018.05.021

[7] P. E. Rapp *et al.*, "Traumatic brain injury detection using electrophysiological methods," *Front. Hum. Neurosci.*, vol. 9, 2015. https://doi.org/10.3389/fnhum.2015.00011

[8] I. Winkler, S. Haufe, and M. Tangermann, "Automatic classification of artifactual ICA-components for artifact removal in EEG signals," *Behav. Brain Funct.*, vol. 7, 2011. https://doi.org/10.1186/1744-9081-7-30

[9] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "ICLabel: An automated electroencephalographic independent component classifier, dataset, and website," *NeuroImage*, vol. 198, pp. 181–197, 2019. https://doi.org/10.1016/j.neuroimage.2019.05.026

[10] M. Klug and K. Gramann, "Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments," *Eur. J. Neurosci.*, vol. 54, no. 12, pp. 8406–8420, 2021. https://doi.org/10.1111/ejn.14992

[11] W. Peterson, N. Ramakrishnan, K. Browder, N. Sanossian, P. Nguyen, and E. Fink, "Differentiating ischemic stroke patients from healthy subjects using a large-scale, retrospective EEG database and machine learning methods," *J. Stroke Cerebrovasc. Dis.*, vol. 33, no. 6, p. 107714, 2024. https://doi.org/10.1016/j.jstrokecerebrovasdis.2024.107714

[12] A. Sawan, M. Awad, R. Qasrawi, and M. Sowan, "Hybrid deep learning and metaheuristic model based stroke diagnosis system using electroencephalogram (EEG)," *Biomed. Signal Process. Control*, vol. 87, p. 105454, 2024. https://doi.org/10.1016/j.bspc.2023.105454

[13] N. Vivaldi, M. Caiola, K. Solarana, and M. Ye, "Evaluating performance of EEG data-driven machine learning for traumatic brain injury classification," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 11, pp. 3205–3216, 2021. https://doi.org/10.1109/TBME.2021.3062502

[14] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, Denver, CO, USA, 1996, pp. 145–151.

[15] F. Lopes *et al.*, "EPIC: Annotated epileptic EEG independent components for artifact reduction," *Sci. Data*, vol. 9, no. 1, p. 512, 2022. https://doi.org/10.1038/s41597-022-01524-x

[16] A. Delorme, "EEG is better left alone," *Sci. Rep.*, vol. 13, no. 1, p. 2372, 2023. https://doi.org/10.1038/s41598-023-27528-0

[17] I. Obeid and J. Picone, "The temple university hospital EEG data corpus," *Front. Neurosci.*, vol. 10, p. 196, 2016. https://doi.org/10.3389/fnins.2016.00196

[18] M. Caiola, A. Babu, and M. Ye, "EEG classification of traumatic brain injury and stroke from a nonspecific population using neural networks," *PLOS Digit. Health*, vol. 2, no. 7, p. e0000282, 2023. https://doi.org/10.1371/journal.pdig.0000282

[19] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn. (ICML)*, Lille, France, 2015, pp. 448–456.

[20] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, Haifa, Israel, 2010, pp. 807–814.

[21] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, 2015, pp. 1–15.

[23] A. Paszke *et al.*, "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, Vancouver, BC, Canada, 2019, pp. 8026–8037.

[24] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006. https://doi.org/10.1016/j.patrec.2005.10.010

[25] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manage.*, vol. 45, no. 4, pp. 427–437, 2009. https://doi.org/10.1016/j.ipm.2009.03.002

[26] L. Xiong *et al.*, "Autonomic dysfunction predicts clinical outcomes after acute ischemic stroke: A prospective observational study," *Stroke*, vol. 49, no. 1, pp. 215–218, 2018. https://doi.org/10.1161/STROKEAHA.117.019312

[27] J. Maruta, M. Suh, S. N. Niogi, P. Mukherjee, and J. Ghajar, "Visual tracking synchronization as a metric for concussion screening," *J. Head Trauma Rehabil.*, vol. 25, no. 4, pp. 293–305, 2010. https://doi.org/10.1097/HTR.0b013e3181e67936

[28] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: A systematic review," *J. Neural Eng*., vol. 16, no. 5, p. 051001, 2019. https://doi.org/10.1088/1741-2552/ab260c

[29] J. A. Urigüen and B. García-Zapirain, "EEG artifact removal—state-of-the-art and guidelines," *J. Neural Eng*., vol. 12, no. 3, p. 031001, 2015. https://doi.org/10.1088/1741-2560/12/3/031001

[30] T. B. Frøsig, "Integrating EEG analysis into game-based learning: A pilot study on memory encoding and learner perceptions in digital educational escape games," *Int. J. Emerg. Technol. Learn. (iJET)*, vol. 20, no. 3, pp. 82–90, 2025. https://doi.org/10.3991/ijet.v20i03.55771

[31] N. M. G. B. da Costa, E. Marçal, R. B. Azevedo, B. L. S. dos Santos, and M. M. de Carvalho, "Evaluating cognitive aspects of ADHD students using brain-computer interface and a digital game: A study in Brazil," *Int. J. Emerg. Technol. Learn. (iJET)*, vol. 20, no. 1, pp. 23–32, 2025. https://doi.org/10.3991/ijet.v20i01.51403

[32] Q. Lu and H. Zhang, "Adaptive learning systems based on deep learning for the diagnosis and support of learning disabilities," *Int. J. Interact. Mob. Technol. (iJIM)*, vol. 19, no. 15, pp. 66–80, 2025. https://doi.org/10.3991/ijim.v19i15.57103

8 AUTHORS

Jialu Gao is a Ph.D. candidate at the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang, Selangor, Malaysia. She received the B.S. degree in computer science from Yunnan University of Finance and Economics, Kunming, China, in 2019 and the Master of Information Technology from the University of Sydney, Sydney, Australia, in 2021. Her research interests include biomedical signal processing, brain–computer interfaces, deep learning, and neurological disease classification using EEG signals (E-mail: gs67746@student.upm.edu.my).

Norwati Mustapha is an Associate Professor with the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang, Selangor, Malaysia. She received the B.S. degree in computer science from Universiti Pertanian Malaysia in 1991, the M.Sc. degree in computing from the University of Leeds, U.K., in 1995, and the Ph.D. degree in artificial intelligence from Universiti Putra Malaysia in 2005. She has authored or co-authored more than 100 publications in international journals and conference proceedings. Her research interests include data mining, machine learning, soft computing, and big data analytics, with a particular focus on applying intelligent methods to medical data analysis and pattern recognition (E-mail: norwati@upm.edu.my).

Razali Bin Yaakob is an Associate Professor with the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang, Selangor, Malaysia. He received the B.S. and M.S. degrees in computer science from Universiti Putra Malaysia in 1995 and 1999, respectively, and the Ph.D. degree in computer science (neural networks) from the University of Nottingham, U.K., in 2008. His research interests include artificial neural networks, deep learning, evolutionary computation, and intelligent systems, with active involvement in applying computational intelligence techniques to healthcare and medical diagnosis (E-mail: razaliy@upm.edu.my).

Noridayu Binti Manshor is a Senior Lecturer with the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang, Malaysia. She received the B.S. and M.S. degrees in computer science from Universiti Putra

Malaysia and Universiti Teknologi Malaysia in 2000 and 2004, respectively, and the Ph.D. degree in computer science (computer vision) from Universiti Sains Malaysia in 2013. Her research interests include computer vision, image processing, pattern recognition, and artificial intelligence, with recent work focusing on feature extraction methodologies and deep learning applications for signal and image classification tasks (E-mail: ayu@upm.edu.my).