Deep Temporal Learning Method for Amyotrophic Lateral Sclerosis Detection on EEG Signals

Sreedevi Chikkudu
Department of Networking and Communications, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu District, Tamil Nadu, India
sc4708@srmist.edu.in

Suresh Annamalai
Department of Networking and Communications, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu District, Tamil Nadu, India
prisu6esh@yahoo.com (corresponding author)

Received: 18 February 2026 | Revised: 21 March 2026, 4 April 2026, and 6 April 2026 | Accepted: 10 April 2026

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | *DOI: https://doi.org/10.48084/etasr.18243*

ABSTRACT

Neurodegenerative disorders constitute a major global contributor to physical disability. In particular, Amyotrophic Lateral Sclerosis (ALS) is a type of neurodegenerative disease that greatly affects life expectancy because it damages nerve cell functions in the Central Nervous System (CNS). Deep Learning (DL) methods have achieved positive results, specifically in the diagnosis of diseases influenced by neurological disorders and in classifying and identifying neurological or psychiatric disorders. This study presents a Deep Temporal Learning Framework for Amyotrophic Lateral Sclerosis Detection (DTLF-ALSD) approach that uses EEG signals. The proposed workflow begins with comprehensive signal preprocessing, including bandpass filtering to remove noise and artifacts, normalization to standardize amplitude distributions, and visual inspection to ensure signal quality enhancement. After preprocessing, three temporal DL architectures, TemporalEEG, TemporalConvNeXtEEG, and TemporalConvNeXtAttEEG, are employed to model dynamic temporal dependencies in EEG signals. These models are designed to extract hierarchical and context-aware representations from multichannel EEG recordings. Subsequently, the learned representations are passed to a fully connected classification head for final decision-making. The models are optimized using RMSprop to ensure stable convergence and improved learning efficiency. The proposed DTLF-ALSD method was examined on the standard EEGET-ALS dataset. Comparative results demonstrate better performance than existing methods across diverse measures for the detection of ALS.

Keywords-amyotrophic lateral sclerosis; electroencephalogram; fully connected; ConvNeXt; attention pooling

I. INTRODUCTION

Amyotrophic Lateral Sclerosis (ALS) is an advanced neurodegenerative disorder, defined by the deterioration of motor neurons in the brain, brainstem, and spinal cord [1]. Early identification of ALS remains a challenging task due to subtle and heterogeneous clinical manifestations [2]. Electroencephalogram (EEG), a non-invasive and relatively inexpensive tool, is widely employed to examine the electrical activity of the brain with improved temporal precision, and can be used to detect ALS effectively [2]. Recent studies have focused on the application of Artificial Intelligence (AI) techniques for ALS detection [3]. Machine Learning (ML) and Deep Learning (DL) techniques are essential for medical disease diagnosis, allowing the automatic learning of hierarchical patterns from complex healthcare data [3]. In particular, DL architectures, namely Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and hybrid models, offer significant results in examining different data sources, such as EEG signals, medical images, and speech recordings for the detection of ALS [4].

In [5], a multi-modal DL structure combined invasive healthcare reports with non-invasive morphology characteristics of Mel-spectrograms extracted from patient speech audio. This model incorporated a multi-modal diagnostic structure, efficiently integrating both data resources to obtain consistent diagnostic precision. This study involved dual experimental states. In the primary state, the audio-trained technique employed a CNN and a systematic optimizer by testing differences in network depth, feature fusion models, and layer dropout likelihoods to enhance the technique's

generalization and stability. In [6], a model was proposed to improve the diagnostic accuracy of ALS, employing ML models on EEG signals. The EEG data from patients with ALS and healthy controls, based on advanced EEG-driven brain signal processing and a sophisticated identification model, were classified using SVM and KNN to identify additional shared normal EEG characteristics through Multiple Sclerosis, ALS, and Parkinson's, called Common Mode Features (CMF), signifying the likelihood of a shared biomarker for these neurodegenerative states. In [7], synthetic EEG signals were created for patients with ALS employing a Conditional Wasserstein Generative Adversarial Network (CWGAN), trained on private EEG data to learn the distribution of ALS EEG signals and create a real synthetic instance. In [8], the role of the Type I Interferon (IFN) pathway in ALS was investigated, evaluating the therapeutic potential of a Type I IFN pathway inhibitor utilizing transcriptomic analysis and a mouse model of ALS.

In [9], VGG19, a CNN-based model, was used to diagnose smaller subdivisions of neurodegenerative disorders, particularly ALS, which decreases the ability of the nervous system to control body muscles. This study utilized a larger collection of EEG records acquired from patients with neurological disorders. In [10], an everyday assistance model was proposed for patients with ALS, based on a multi-modal wearable mouse. This method included two submodels: a wheelchair system that depended on a mouse system to support the lower and upper boundaries, and a BCI headband, where ALS patients could manage a computer cursor on the screen with eye blinking and slight head rotation, and more computer options to drive a wheelchair with particularly intended Graphical User Interfaces (GUIs). In [11], an innovative 3D deep CNN (3D-CNN) technique, named MEGNet3D, was suggested to distinguish between ALS and healthy people from their resting condition sensor-level Magnetoencephalography (MEG) information. The raw MEG information was transformed into its time-frequency representation, which was then used as input to MEGNet3D. In [12], the features of sleep diseases were investigated, along with their relation to abnormal white-matter integrity of patients with sporadic ALS. Sleep diseases were identified by employing the Epworth sleepiness scale, polysomnography, and Pittsburgh sleep quality index; patients were identified as those with good and poor sleep quality. The relations among scores in impaired white-matter tracts and the Pittsburgh Sleep Quality Index were analyzed by employing multiple regression.

In [13], an automatic Intensive Neural Classifier Model based on Deep Learning (DL-INC) was designed to support medical experts in the monitoring of ALS. As the EEG signals are non-linear and oscillatory, this DL-INC technique inspects the signal by employing the diverse frequency sub-bands. In [14], an enhanced algorithmic architecture for EEG characterization was presented to explore the medical condition of ALS patients. In [15], a transformer-based technique was proposed to identify ALS by employing the EEG and eye-tracking dataset of ALS patients (EEGET-ALS), which includes a total of 1,989 data instances. This method indicates a major development in the field of neurodegenerative disease diagnosis, possibly enhancing patient outcomes and quality of life by applying a two-minute recording.

This study presents a Deep Temporal Learning Framework for Amyotrophic Lateral Sclerosis Detection (DTLF-ALSD) using EEG signals. The main objective is to examine how hybrid temporal convolutional and attention-based architectures enhance ALS classification performance from EEG signals. The workflow integrates dataset balancing, signal preprocessing, structured feature generation, and three temporal DL architectures. A Fully Connected (FC) classification head is optimized using RMSprop to ensure stable convergence and efficient learning. The experimental results confirm the robustness and improved predictive accuracy of the proposed approach.

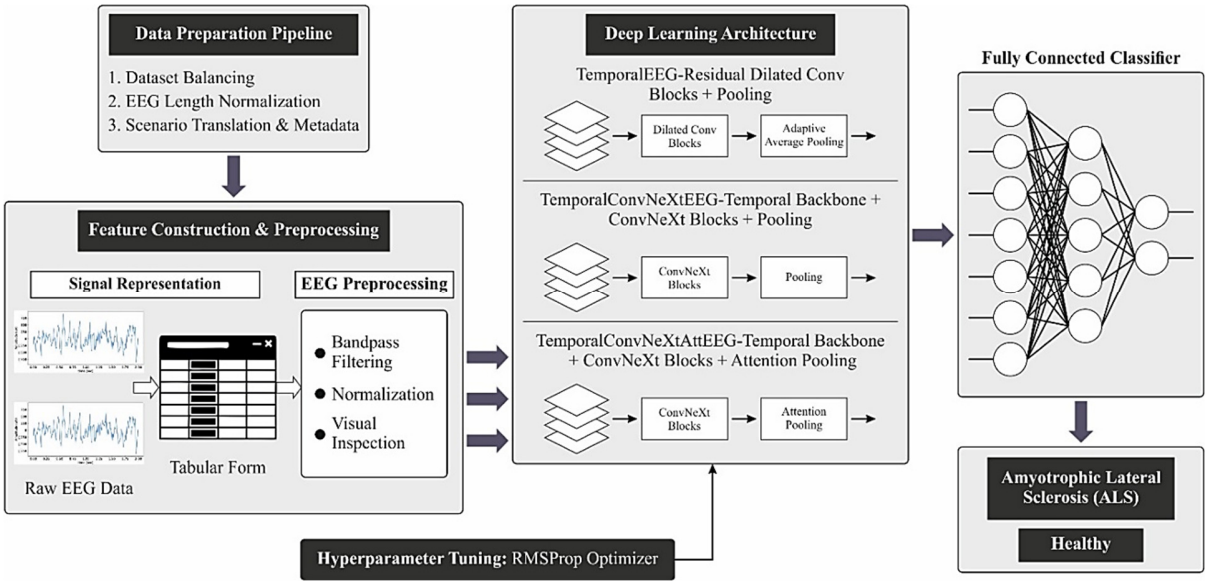

Fig. 1. Overall flow of DTLF-ALSD framework.

II. METHODOLOGICAL FRAMEWORK

The proposed DTLF-ALSD framework leverages EEG signals to enhance classification performance. Figure 1 exemplifies the entire flow of the proposed DTLF-ALSD methodology. The framework involves a single preprocessing step, the three temporal DL architectures, a classification head, and an optimizer.

A. Data Used

The effectiveness of the proposed technique was thoroughly validated using the EEGET-ALS dataset [16], which comprises the EEG and eye tracking recording data gained from six patients with ALS in a locked-in state and 170 healthy individuals. The dataset encompasses two classes, namely ALS and Healthy. To ensure consistent model training, the raw EEG dataset is preprocessed in a structural pipeline. Primarily, class imbalance is solved to quantify the number of EEG files for healthy participants and those with ALS. Afterward, the EEG length is normalized to achieve standardization of the temporal dimension of the input.

From the recordings, fewer than 4096 instances are removed to uphold signal integrity, whereas larger recordings are shortened to a fixed length of 4096 instances. At last, scenario-level metadata collected in JSON format is translated into English and systematically analyzed for extracting class labels such as Healthy or ALS, file path references, and scenario identifiers.

B. Signal Representation and Preprocessing

All EDF files cover multichannel EEG signals. A Butterworth bandpass filter is used to improve the EEG signals prior to classification and feature extraction. This filter is chosen because of the maximum flat frequency response from the passband and minimum signal distortion, which makes it particularly appropriate for EEG signal processing. Z-score normalization is frequently used in EEG data processing to standardize the information, resulting in a mean of zero and a standard deviation of one, normalizing the use of all information in the databases, ensuring that every feature has a similar scale. Visual analysis of waveform comparison plots was used to validate the effectiveness of signal enhancement models. Such plots illustrate the processed and combined real signals, allowing for a direct evaluation of improvements in noise reduction, feature preservation, and amplitude clarity. In comparing waveform patterns, this step can highlight how the enhancement technique overwhelms undesirable alterations while maintaining important signal features. The constant transitions and increased peak visibility in the enhanced waveform confirm the efficacy of the utilized processing model. Consequently, the waveform comparison plots provide an instinctive and qualitative validation of signal enhancement outcomes.

C. DL Model Architectures

To assess the efficiency of temporal representation learning for ALS classification, three gradually improved temporal structures were applied. In each method, the network is conceptually separated into dual major elements: (i) a temporal feature extraction backbone and (ii) an FC classification head.

1) TemporalEEG (Baseline Model)

An enhanced Residual Dilated Convolution (RDC) model was used to improve the network's feature extraction ability. This model takes feature data over a wider range to introduce convolutional layers with various dilation rates and a feature fusion module, enhancing the capability of capturing detailed features and intricate scenes. The input feature mapping Y is converted across a sequence of convolutional layers with several dilation rates, ensuring that multi-scale characteristics are effectively captured and integrated. The enhanced RDC model contains four branches, with convolutional dilation values of 1, 3, and 5. These branches are intended for capturing features at various scales. The convolution operation is given by:

$$F_i(Y) = \sigma(W_i * Y + b_i) \quad (1)$$

where W_i denotes the weights of the convolution layer and b_i signifies the bias. In the RDC model, a 1×1 convolution is employed for reducing dimensions and capturing local features. The 3×3 convolution dilation level of 1 obtains typical convolutional parts. The 3×3 convolution dilation rate of 3 captures middle-range context to expand the receptive area. The 3×3 convolutional dilation rate of 5 obtains larger contextual data. After processing throughout various branches of dilated convolution, each branch's outcome is fused and concatenated with a 1×1 convolutional layer as follows:

$$F(Y) = \sigma(W_f * concat\left(F_1(Y), F_2(Y), F_3(Y), F_4(Y) + b_f\right) \quad (2)$$

Let b_f and W_f represent the biases and weights of a 1×1 convolutional state. The outcome is achieved by a residual connection to add the input mapped feature processes by the 1×1 convolutional layer to the outcome of the enhanced RDC model, as given in:

$$I = Y'' + F(Y) \quad (3)$$

where Y'' denotes the outcome managed throughout the 1×1 convolutional layer. The feature mapping is handled with a dilated residual model, maintaining the real characteristics while improving the multi-scale context.

This study regulates the fusion method and convolutional kernel dimension by reducing the problem of local detail retention deficiency and enhanced computational difficulty affected by higher dilated rates. Spatially divisible convolutions substitute a 5×5 module with dual 5×1 and 1×5 convolution models, upholding the receptive region but decreasing network limits and increasing non-linearity and technique expression. A 1×1 convolutional layer is inserted after all branches for adjusting the dimension and linear excitation, decreasing computation and parameters, and improving efficacy. Global Average Pooling (GAP) creates feature layers that are more simply changed into classifier probabilities, evading overfitting.

$$GAP = \frac{1}{H \times M} \sum_{h=1}^{H} \sum_{m=1}^{M} Y_{ab} \quad (4)$$

where H and M define the height and width of feature mapping, $Y_a b$ denotes the value of pixels at column b and row

a in the feature mapping, $\sum_{h=1}^{H}\sum_{m=1}^{M} Y_{ab}$ signifies the overall value of pixels in the feature mapping, and $\frac{1}{H\times M}$ implies the normalization factor.

2) TemporalConvNeXtEEG (Hybrid Model)

To improve representational capability, the TemporalConvNeXtEEG architecture extends the baseline by integrating the ConvNeXt-assisted residual blocks next to the temporal network. The temporal backbone helps as the primary feature extractor. This is processed to the multichannel EEG signals for gaining the short-term temporal dependency and primary spatial-temporal patterns. By using the stacked temporal convolutional layers, the backbone converts original EEG data into useful transitional feature maps. This phase concentrates on learning lower-level transient patterns related to neurological settings and temporal patterns such as rhythmic oscillations.

After the backbone, the ConvNeXt contains numerous ConvNeXt blocks. The ConvNeXt [17] updates the CNN framework by presenting a substantial aspect, efficiently integrating the robustness of CNNs with the benefits of the transformer model. The ConvNeXt architecture employs a hierarchical measure of convolutional blocks, which could be continuously expanded to reproduce particular features of the ViT architecture. However, it is important to note that ConvNeXt retains the convolutional processes and does not utilize the Multi-Head self-attention mechanisms (MHA) determined in the transformer model. The ConvNext module is employed for Depth-Wise Convolution (DWC) 7×7 dimensions, and then the parameter y LayerNorm with a 1×1 DWS and four times as many channels as the actual convolutions. Subsequently, it is decreased by a GELU activation and lastly a 1×1 DWC with a similar count of channels primarily. This enables ConvNeXt to learn a highly wide-ranging representation of the EEG, offering improved evaluation performance. Although ConvNeXt surpasses other methods in gaining the local spatial features and proficiently removing the features from EEG, it mainly depends on convolutional processes that can limit its capability for acquiring the global correlations and long-term dependency in the data. This integration of normalization, Depth-Wise Separable Convolutions (DWSC), and the MLP-based channel integration allows effective learning while maintaining spatial locality.

The TemporalConvNeXtAttEEG framework highly extends the hybrid method by presenting a Temporal Attention and Pooling (TAP) mechanism. This is a DL method that integrates the sequential features by transferring various significance weights for every individual time step. In contrast with max or average pooling, whole temporal features are similarly chosen, or only the strongest activation, and the TAP is learned to focus on more descriptive parts of the sequences.

In the processing of the sequential data, the whole-time steps are not contributed similarly to the last prediction. During EEG analysis, particular temporal segments comprise discriminatory intellectual patterns, whereas alternatives are added to noise. The TAP learns adaptable weights to focus on useful and suppress the unrelated segments. Assuming the form $H = \{h_1, h_2, \ldots, h_T\}$, attention scores e_t are calculated by applying learning parameters and normalization corresponding to softmax for gaining the weights α_t. The last mathematical expression is given by:

$$z = \sum_{t=1}^{T} \alpha_t h_t \quad (5)$$

where α_t indicates the significance of every individual time step.

3) Classifier: FC Layer

The FC layer involves a compact architecture with a 6-6-8-2 topology, comprising an input layer with 6 neurons, two hidden layers with 6 and 8 neurons correspondingly, and an output layer with 2 neurons equivalent to the classification classes (ALS and Healthy). This streamlined representation eliminates redundancy while clearly conveying the model structure. The FC layers are the most prevalent in NNs and are utilized for most frameworks. Every individual node is interlinked to every node from the preceding and succeeding layers. The main operation of an FC layer is to convert the feature space to make it more flexible. With this modification, size is increased, reduced, or unchanged. In every context, the novel dimensions have the linear integration of dimensions in the preceding layer. By employing an activation function, the dimensions are obtained as non-linear. FC layers are created for all types of interconnections among the potential input parameters. As a result of this architectural agnostic method based on adequate depth and width, FC layers possess the intellectual capacity for learning all functions.

Although real-world involvement is demonstrated, conceptual ability is not generally accomplished. Researchers have encountered these difficulties by emerging more specific layers, namely, recurrent and convolutional layers. These layers have essentially taken the benefits of the preliminary bias with respect to sequential or spatial models of particular categories of data. This method can improve classification performance by presenting significant bias to FC layers via the incorporation of intrinsic data existing in EEG instances.

RMSProp [18] is an optimizer that can be considered as a basic version of the Adam method, ignoring the momentum factor. This is significant for its robust empirical outcomes, which frequently surpass those of the same methods, namely Stochastic Gradient Descent (SGD).

$$v(t) = \beta v(t-1) + (1-\beta)\left(\frac{\partial l}{\partial w}\right)^2 \quad (6)$$

$$w(t) = w(t-1) - \frac{\eta}{\sqrt{v(t)}}\frac{\partial l}{\partial w} \quad (7)$$

Here, v signifies the moving average of squared gradients, whereas $\partial c/\partial w$ represents the gradient of the cost function with respect to weights. The constraint η denotes the rate of learning, and β denotes the moving restriction of the average.

III. EXPERIMENTAL ANALYSIS

The proposed model was developed using Python 3.6.5 on a PC with an i5-8600k, GeForce 1050Ti 4 GB, 16 GB RAM, 250 GB SSD, and 1 TB HDD. The experimental analysis of the DTLF-ALSD method was examined using the EEGET-ALS

dataset [16], which comprises recordings from ALS and Healthy persons, forming a binary classification problem (2 class labels). For experimental validation, the dataset was split into 70% for training and 30% for testing. Figure 2 illustrates the preprocessing procedure.

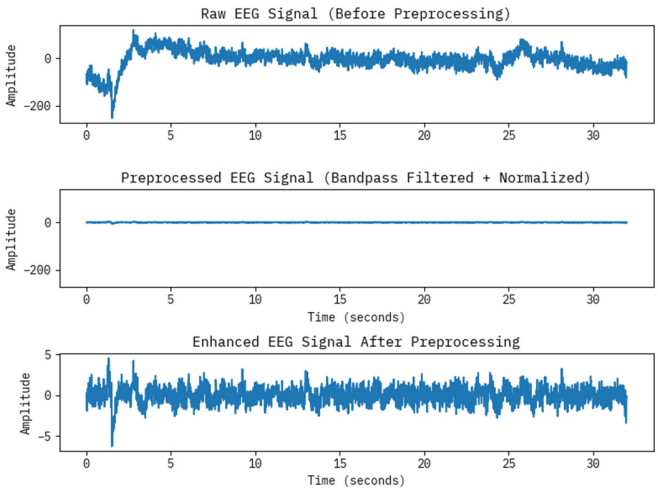

Fig. 2. Preprocessing.

Figure 3 represents the confusion matrices for the training and validation sets for TEEG, TCEEG, and TCAEEG in ALS and Healthy conditions. The simulation outcome values specify that the TEEG, TCEEG, and TCAEEG models excellently detect and precisely classify two classes.

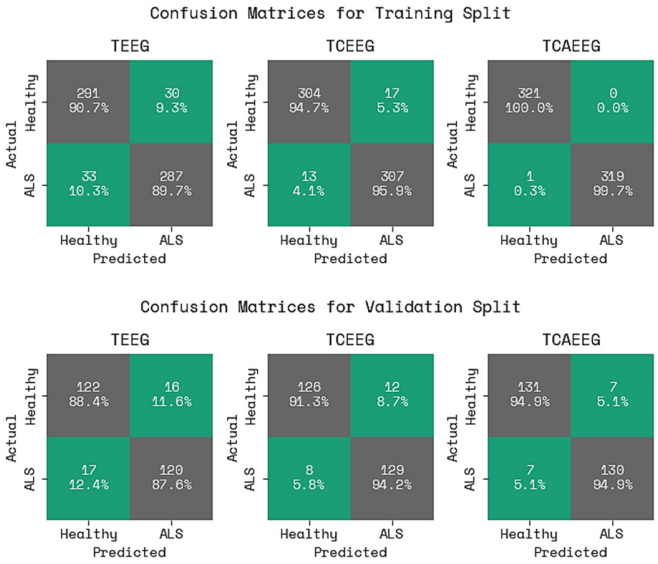

Fig. 3. Confusion matrices of TEEG, TCEEG, and TCAEEG models with training and validation splits.

Table I shows a comparison of the DTLF-ALSD approach with existing models [6, 19, 20]. The baseline models portrayed in Table I were reimplemented based on their original descriptions and evaluated on the EEGET-ALS dataset using a unified preprocessing pipeline and consistent evaluation protocol instead of being directly adopted from previously published results. The results demonstrate that the RNN approach had lower performance, whereas the HNN, GRU, and SVM methods were slightly better. The KNN, LDA, and WebNet methods displayed significant performance improvements, whereas the DTLF-ALSD (TEEG) approach realized superior results with $accur_y$ of 95.32%, $preci_n$ of 94.75%, $sensi_y$ of 95.94%, $speci_y$ of 94.70%, and $F1$ of 95.34%.

TABLE I. COMPARATIVE ANALYSIS OF DTLF-ALSD METHOD WITH EXISTING MODELS

Methods	$Accur_y$	$Preci_n$	$Sensi_y$	$Speci_y$	F1
SVM [6]	87.00	88.36	87.16	85.53	93.11
KNN [6]	91.74	91.44	85.78	86.85	90.04
LDA [19]	90.00	86.93	86.85	86.88	89.61
HNN [19]	82.06	90.96	93.99	87.49	89.87
RNN [19]	78.19	93.10	88.16	89.83	88.17
WebNet [20]	90.26	91.38	87.13	87.07	90.47
GRU [20]	86.57	93.09	85.31	93.56	90.71
DTLF-ALSD (TEEG)	95.32	94.75	95.94	94.70	95.34
DTLF-ALSD (TCEEG)	95.32	94.75	95.94	94.70	95.34
DTLF-ALSD (TCAEEG)	99.84	100.00	99.69	100.00	99.84

The models from previous studies were reimplemented using a unified preprocessing pipeline.

The DTLF-ALSD (TCEEG) approach achieved even better performance with $accur_y$ of 95.32%, $preci_n$ of 94.75%, $sensi_y$ of 95.94%, $speci_y$ of 94.70%, and $F1$ of 95.34%. The DTLF-ALSD (TCAEEG) model presents outstanding performance with $accur_y$ of 99.84%, $preci_n$ of 100.00%, $sensi_y$ of 99.69%, $speci_y$ of 100.00%, and $F1$ of 99.84%. These results confirm that the DTLF-ALSD method realizes better learning performance compared to the other techniques.

IV. CONCLUSION

This study presented a novel DTLF-ALSD model by using EEG signals. The primary objective was to explore how hybrid temporal convolutional and attention-based frameworks enhance ALS classification performance from EEG signals. Initially, the signal preprocessing phase comprises bandpass filtering to remove noise, normalization to regulate amplitude distributions, and visual inspection to ensure improved signal quality. After preprocessing, three temporal DL methods, namely TemporalEEG, TemporalConvNeXtEEG, and TemporalConvNeXtAttEEG, were exploited to model dynamic temporal dependencies in EEG signals. These approaches were intended to extract hierarchical and context-aware representations from multichannel EEG data. Consequently, the learned representations were passed to an FC classification head for final decision-making. The models were optimized utilizing RMSprop to ensure stable convergence and enhanced learning efficiency. The proposed DTLF-ALSD approach was experimentally investigated on the benchmark EEGET-ALS dataset. The comparative results exhibited better performance over existing models with respect to various aspects, leading to improved classification performance for ALS detection.

DECLARATION OF COMPETING INTERESTS

The authors declare no conflict of interest.

ACKNOWLEDGEMENT

Not applicable in this work.

DATA AVAILABILITY STATEMENT

The dataset used to train and test the proposed approach is publicly available at [16].

REFERENCES

[1] J. Morris, "Amyotrophic Lateral Sclerosis (ALS) and Related Motor Neuron Diseases: An Overview," *The Neurodiagnostic Journal*, vol. 55, no. 3, pp. 180–194, Sept. 2015, https://doi.org/10.1080/21646821.2015.1075181.

[2] H. Uyanik, A. Sengur, M. Salvi, R. S. Tan, J. H. Tan, and U. R. Acharya, "Automated Detection of Neurological and Mental Health Disorders Using EEG Signals and Artificial Intelligence: A Systematic Review," *WIREs Data Mining and Knowledge Discovery*, vol. 15, no. 1, 2025, Art. no. e70002, https://doi.org/10.1002/widm.70002.

[3] A. Alghamdi, T. Alsubait, A. Baz, and H. Alhakami, "Healthcare Analytics: A Comprehensive Review," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6650–6655, Feb. 2021, https://doi.org/10.48084/etasr.3965.

[4] A. Dixit, V. Bajaj, and P. K. Padhy, "ALSNet: A Lightweight Deep Learning Framework for ALS Detection Using Enhanced TF Images From EMG Sensor Data," *IEEE Sensors Journal*, vol. 26, no. 5, pp. 7433–7441, Mar. 2026, https://doi.org/10.1109/JSEN.2026.3655954.

[5] I. N. Switrayana, T. T. Sujaka, and I. S. Putri, "A Multimodal Deep Learning Framework for Amyotrophic Lateral Sclerosis Diagnosis using Clinical and Audio Morphology Features," *Sistemasi: Jurnal Sistem Informasi*, vol. 15, no. 1, pp. 220–236, Jan. 2026, https://doi.org/10.32520/stmsi.v15i1.5763.

[6] A. Abedi, G. Moradi, R. S. Shirazi, and M. Jahed, "Improving ALS Diagnosis Based on Human EEG Signal Analysis with Machine Learning," in *2024 31st National and 9th International Iranian Conference on Biomedical Engineering (ICBME)*, Nov. 2024, pp. 63–68, https://doi.org/10.1109/ICBME64381.2024.10895722.

[7] A. Mutlu, Ş. Doğan, and T. Tuncer, "Synthetic ALS-EEG Data Augmentation for ALS Diagnosis Using Conditional WGAN with Weight Clipping." arXiv, June 19, 2025, https://doi.org/10.48550/arXiv.2506.16243.

[8] O. Carletta *et al.*, "Genotype-specific interferon signatures in amyotrophic lateral sclerosis relate to disease severity," *Brain*, vol. 149, no. 2, pp. 489–501, Feb. 2026, https://doi.org/10.1093/brain/awaf324.

[9] N. Lohia, C. Mathew, and G. Shankel, "Applying Transfer Learning and Existing EEG Datasets to Identify Patients With ALS," *SMU Data Science Review*, vol. 8, no. 2, Sept. 2024.

[10] Y. Jiang, K. Li, Y. Liang, D. Chen, M. Tan, and Y. Li, "Daily Assistance for Amyotrophic Lateral Sclerosis Patients Based on a Wearable Multimodal Brain-Computer Interface Mouse," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 33, pp. 150–161, 2025, https://doi.org/10.1109/TNSRE.2024.3520984.

[11] K. Samanta *et al.*, "An Automated Detection of Amyotrophic Lateral Sclerosis from Resting State MEG Data Using 3D Deep Convolutional Neural Network," *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 3337–3342, Jan. 2024, https://doi.org/10.1109/SMC53992.2023.10393987.

[12] X. Li *et al.*, "Sleep disorders and white matter integrity in patients with sporadic amyotrophic lateral sclerosis," *Sleep Medicine*, vol. 109, pp. 170–180, Sept. 2023, https://doi.org/10.1016/j.sleep.2023.07.003.

[13] R. B. Revathi and T. P. Ramachandran, "A Novel Deep Learning Model to Predict Amyotrophic Lateral Sclerosis (ALS) Disease Based on Intensive Neural Classifiers," in *2025 IEEE 2nd International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS)*, Aug. 2025, pp. 1–7, https://doi.org/10.1109/ICITEICS64870.2025.11341718.

[14] Y. Zheng, K. Li, and Y. Jiang, "Decoding Neural Mechanisms of Meditation in Amyotrophic Lateral Sclerosis Patients via Multichannel Electroencephalogram Time-Frequency Analysis," in *2025 2nd International Conference on Electronic Engineering and Information Systems (EEISS)*, May 2025, pp. 1–6, https://doi.org/10.1109/EEISS65394.2025.11085618.

[15] S. Jain, "A Transformer-Based Approach to Diagnose Amyotrophic Lateral Sclerosis via Electroencephalogram Analysis," in *2024 17th International Conference on Advanced Computer Theory and Engineering (ICACTE)*, Sept. 2024, pp. 88–93, https://doi.org/10.1109/ICACTE62428.2024.10871346.

[16] "EEGET-ALS Dataset." figshare, June 20, 2024, https://doi.org/10.6084/m9.figshare.24485689.v1.

[17] Z. Liu, H. Mao, C. Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 11966–11976, https://doi.org/10.1109/CVPR52688.2022.01167.

[18] T. Tieleman and G. Hinton, "Lecture 6.5-rmsprop Divide the Gradient by a Running Average of Its Recent Magnitude.," *Coursera Neural Networks for Machine Learning*, 2012.

[19] P. Zych, K. Filipek, A. Mrozek-Czajkowska, and P. Kuwałek, "Classification of Electroencephalography Motor Execution Signals Using a Hybrid Neural Network Based on Instantaneous Frequency and Amplitude Obtained via Empirical Wavelet Transform," *Sensors*, vol. 25, no. 11, May 2025, https://doi.org/10.3390/s25113284.

[20] H. Xu, W. Haider, M. Z. Aziz, Y. Sun, and X. Yu, "Transforming Motor Imagery Analysis: A Novel EEG Classification Framework Using AtSiftNet Method," *Sensors*, vol. 24, no. 19, Oct. 2024, https://doi.org/10.3390/s24196466.