International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Mathematical Modelling of Engineering Problems
Vol. 12, No. 3, March, 2025, pp. 860-874
Journal homepage: http://iieta.org/journals/mmep

Classification of EEG Signal Using Deep Neural Network Techniques

Hanan Badeea Ahmed[1], Mohammed Sami Mohammed[2], Adham Hadi Saleh[1*]

[1] Department of Electronic Engineering, College of Engineering, University of Diyala, Baqubah 32001, Iraq
[2] Department of Computer, College of Education for Pure Science, University of Diyala, Baqubah 32001, Iraq

Corresponding Author Email: adham.hadi@yahoo.com

Copyright: ©2025 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/mmep.120312

Received: 5 August 2024
Revised: 25 November 2024
Accepted: 2 December 2024
Available online: 31 March 2025

Keywords:
LSTM, EEG signal, feature selection, correlation calculator

ABSTRACT

The robustness and accuracy of EEG signal categorization have significantly increased as a result of recent developments in deep learning. By combining correlation-based feature selection with Long Short-Term Memory (LSTM) networks, this work presents a novel method that improves the categorization of EEG data through a hybrid model. The suggested LSTM-correlation model outperformed a number of neural network architectures, including traditional machine learning and other deep learning methods, according to thorough comparison, especially while managing intricate real world EEG data. With 97.99% classification accuracy after just two epochs and 73 iterations, 98.99% after five epochs and 306 iterations and 99.98% after 1000 epochs and 5800 iterations, the model demonstrated remarkable classification accuracy. These results highlight how well the model can handle big datasets and its efficiency was demonstrated by training times that range from seconds to minutes. By identifying pertinent characteristics from the EEG data and improving the model performance throughout several training stages, the hybrid LSTM-correlation technique demonstrated remarkable adaptability. When paired with feature selection strategies, this work demonstrates the versatility of LSTM networks and their promise for high-performance jobs requiring sequential data. Furthermore, across various epochs and iterations, the suggested method beat alternative topologies, including Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN), confirming its position as a top EEG classification method. These findings have important practical ramifications, especially for real-time medical diagnostic applications where accurate and effective EEG signal classification is essential. All experiments were conducted using MATLAB R2022b.

1. INTRODUCTION

Since the first electrical recording for EEG signals that belongs to rabbits and monkeys, these signals have been more investigated by authors and researchers. The textual evidence was originally rather noisy and the findings were not particularly strong. Numerous factors such as biofouling, environmental contamination and misaligned electrodes can cause noise in EEG readings. The quality of EEG recordings has improved over several previous years due to developments in signal processing technology and recording apparatus. Following the development of the ECG in the 1930s, EEGs became widely used in medicine and were recorded in new methods. EEG signal recording and analysis are now standard procedures in clinical and medical research. Additionally, EEG has proven to be a helpful tool in many scientific fields, such as human factors, neuropsychology and psychology as explained by Panteliadis [1]. The study of EEG signals and their application in identifying aberrant/normal brain activity is the main topic of this article. It is a physiological technique for capturing brain electrical activity. Despite new methods and technical advancements, EEG remains the most used technology for brain imaging, maintaining its significance. An excellent source of anatomical information is not the EEG, but when analyzed by an experienced EEG professional it offers a plethora of physiological data. While age affects normal EEG patterns, an aberrant EEG signal may point to the existence of a brain illness. Working on EEG, particular clinical indications or symptoms, such as altered mental status, periods of inactivity or involuntary episodes, are identified and compared with the patient's prior EEG and/or the EEG of other patients. Those patients have the same brain problem in order to determine the precise condition. Finding the clinical characteristics of the EEG and the outcomes of other neurological testing is the first stage in the interpretation process. With this knowledge, abnormal EEG data are simple to uncover, recognize and connect to the patient's clinical issue. The study of EEG signal analysis is a large field. In such works, details the processing of EEG signals and its application to autonomous brain structure recognition and deep learning. Research in the field of EEG signal analysis requires a foundational understanding of EEG signals and recording methods as more explained by Arora and Mishra [2]. While the idea was widely established in the 1980s, there has generally been little success in establishing a practical model with suitable performance in a data-rich sector. Between then and 2000, this had a detrimental effect on research capability and investments in the aim of leveraging deep knowledge for artificial intelligence as explained by Galván and Mooney [3], and Adate et al. [4]. The biological intelligence system employs intrinsic structures for decision-making and behavior control according to research in neuroscience science. It was utilized to forecast the effects of their choices and select the optimal courses of solutions in accordance with their objectives. These forecasts are frequently the result of long-term learning and are mostly based on experience. Machine learning (ML) needs to represent these models and deep learning (DL) complies by using several networks to distinguish between different network architectures with expressing the links between inputs and outputs as more clarified by Bikku [5]. It's critical to differentiate between these channels and nerves by emphasizing several levels as opposed to depth. The massive volume of digital data generated and gathered today in addition to the advancements in DL technologies, have got the interesting of researchers as in the study conducted by Ding et al. [6]. Deep Neural Network (DNN) is an incredibly powerful tool for problem solving in domains including natural language processing, computer vision and gaming. DL extracts and transforms data using many layered linear processing techniques. The output of the preceding input stage is utilized in every successful stage. The revolutionary success of DL in a variety of applications can be attributed to their simplicity and uniqueness [7, 8]. DNN deals with techniques that enable machines to learn to execute activities that often require human intelligence. ML is now needed to do complicated activities due to the growth in infrastructure capacity and the intricacy of decision-making processes that mimic human behavior, like brain simulation. This is the point at which much deeper learning becomes relevant and CNN algorithms fail. DL has the primary benefit of optimizing the performance of all feasible methods in contrast to artificial processing, which makes use of a collection of low-level objects. What sets deep learning apart from other forms of machine learning is its feature hierarchy as presented clearly by Xiao et al. [9], and Mohammed and Mohammed [10]. An essential or indirect way to measure brain activity with temporal resolution is the EEG. Electrodes applied to the scalp can be used to record the electrical activity generated by brain neurons. EEG sensors are a popular neuroimaging technique that can be employed in a range of settings. These settings are from clinical to research and from healthy to pathological circumstances, due to its ubiquity and noninvasive nature. Regarding to the distinct characteristics of EEG signals which are intrinsically complex and unstable, ML is necessary in order to identify significant aspects and categorize the signals. To prove reliable output, it is crucial to combine the extraction process with categorization. Brain Computed Interface (BCI) aims to classify brain activity, which is an important way to study brain function. It offers a method for precisely evaluating brain activity, recognizing alterations in brain activity and diagnosing anomalies in brain activity. The diagnosis of epilepsy is a prime illustration of this. On an EEG, high voltage spikes and spikes (HVTS) can be observed. The presence, location, type, and severity of HVTS epilepsy can be identified, which is useful information for clinical diagnosis and anticonvulsant selection. Additional developments in the EEG class can be found in a variety of BCI tasks, including data processing for visuals, thoughts/words, motor imagery and other purposes as more suggested by Saeidi et al. [11], and Hosseini et al. [12]. Figure 1 shows a simple basic steps of EEG signal classifications based on hybrid techniques as explained by Tosun and Çetin [13].

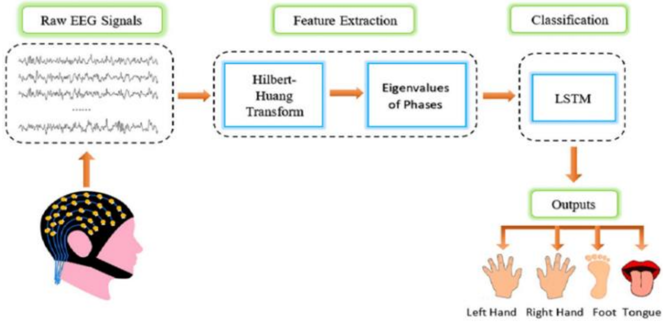


Figure 1. EEG signal classifications based on hybrid techniques

Merlin Praveena et al. [14] employed various DL techniques and diversified architecture in the analysis of EEG signals, providing insight into how to advance AI-based systems to a new level. In addition to being beneficial for those who are investigating EEG signals using DL algorithms, this study was included details on the application of deep learning techniques in EEG signals as well as the difficulties and constraints associated with each technique in categorization. While Guerrero et al. [15] used typical classification techniques in their work for extracting features using Fourier analysis and taking frequency bands into consideration. The data used for these techniques included frequency information in the channels containing distinguishing information from EEG exams. Artificial Neural Networks (ANN) with an accuracy of 86%, were found to be the most effective classification strategy for characterizing individuals with epilepsy in this paper. CNN were employed also by Guo et al. [16] to classify motor imagery EEG data. Authors built an input pipeline for a customized CNN using Spatial Filtering and Short-Time Fourier Transform (STFT). They employed a two-stage CNN to lower depth-to-prediction in order to reduce the numerous parameters. While the second stage of the CNN learns temporal filters, the first stage learns spatial filters. With a basic bespoke model, their accuracy was about 86.3% and when they enhanced a well-known EEG model it was improved to about 89.3%. The literature currently available on EEG-based depression identification identified a number of short comings in the techniques used. Modern methods frequently depend on datasets that are small in size and difficult to acquire, which might introduce biases and limited generalizability. Numerous research concentrated on examining a single dataset, which was not fully convey the intricacy of EEG patterns linked to melancholy. Furthermore, some approaches used exponentially sophisticated deep learning architectures that were computationally demanding, which limited their applicability in real-world situations as mentioned by Al Fahoum [17]. It was interesting to note that although time-frequency techniques had been extensively used to identify aberrant and feature waves in EEG signals, they also had built-in drawbacks that limit their use in EEG analysis which more described by Qin and Ji [18]. Both wavelet transforms and the conventional STFT had advantages and disadvantages indicating the need for more flexible and adaptive methods. A novel integrated methodology that combines deep neural networks and phase space reconstruction had been developed to fill these gaps as demonstrated by Al Fahoum [17]. By employing publicly accessible EEG datasets to reduce biases, reconstructed phase space analysis to better capture intricate EEG patterns and deep neural networks for effective and precise classification, this strategy searched to overcome the drawbacks of existing techniques.

The suggested approach aimed to improve depression detection accuracy while laying the groundwork for expandable, easily accessible mental health solutions that may be used in practical contexts which is also mentioned by Al Fahoum [17]. Furthermore, multi-resolution time-frequency analysis techniques based on wavelet packet transform and STFT had been introduced in virtual EEG recording and analysis equipment with the goal of enhancing signals capacity for self-adaptation and facilitating the more effective identification of fundamental rhythms in EEG signals which also been described in [18]. The complexity of brain activity was difficult for current EEG analysis techniques to fully capture due to a number of limitations. For example, the nonlinear and nonstationary nature of brain waves had not been adequately described by traditional EEG analysis methods as suggested by Cataldo et al. [19] and Ke et al. [20]. Numerous current approaches depended on linearity or orthogonality assumptions that were not accurately represent the physiological processes as introduced by Koenig et al. [21]. Furthermore, conventional methods frequently called for a great deal of de-noising of EEG data, which eliminated crucial signal components as mentioned by Ke et al. [20]. Researchers were investigating more sophisticated methods to overcome some constraints like: non-linear dynamical approaches, notably entropy-based techniques like Multiscale Fuzzy Entropy (MFE) had demonstrated promise in better capturing the complex dynamics of the brain which was also mentioned by Cataldo et al. [19]. Without needing a great deal of preprocessing, these methods highlighted minute distinctions between diseased and healthy brain states. The goal of new modeling approaches was to increase source localization accuracy. For example, to facilitate more accurate EEG forward modeling, particularly in thin cortical structures, current preserving dipolar source models had been created as mentioned by Miinalainen et al. [22]. This strategy suggested certain current techniques by striking a balance between accuracy and performances. To take use of complementary information and attain greater spatiotemporal resolution, multimodal integration of EEG with other neuroimaging methods, including fMRI, was being interested by Kalus et al. [23] and Lei et al. [24]. It had been demonstrated that employing new Bayesian techniques to combine EEG and fMRI data improved the sensitivity of activation detection. This field suggestions were shifting toward more adaptable, data-driven methods that preserve mathematical objectivity and rigor while better capturing the physiological complexity of brain activity.

2. TECHNIQUES FOR MEASURING NEURAL ACTIVITIES IN EEG

Muscle activity can be measured using two primary techniques: invasive and non-invasive. The first method is not taken into consideration in this work because EEG is non-invasive. Determining the location and temporal course of activity is the goal of using EEG to measure brain activity. With the EEG recording head, the source of cerebral activity can now be reliably identified. The EEG signal represents the activity's temporal course. But, assuming that the conduction of electrical activity in the brain is isotropic, homogenous and infinite with the conduction of the skull, it manifestly demonstrates that the source of origin is very little. As a matter of fact, the skull is neither one of its parts nor a competent leader. The EEG signal recorded on the scalp is distorted by the skull, particularly the closely spaced radio dipole components [25]. Techniques for these purposes can be listed as follows:

2.1 Electrode placement

While the number and kind of electrodes needed for various EEG recordings varies, common areas are taken into consideration for particular tests (e.g., 10–20 systems in polysomnography) as shown in Figure 2 [26]. The '10-20' system is a technique that is crucial for long-term research and monitoring because it primarily permits the same electrode to be applied to multiple subjects or patients. It's crucial to take into account how using various electrodes may impact the recording of brain activity. The most common explanation is ear alignment, which may not always be true because studies have shown that this can result in asymmetrical noises that represent brain activity (asymmetrical activity in ADD). Simulating the EEG as soon as possible is vital because of the brain's rapid electrical activity. Whether the electrodes are attached to the head or not, a balance must be struck between the quantity of electrodes employed and the caliber of the output. Although more sample points might yield more information regarding the condition of the brain, recent computational techniques have demonstrated that the entire EEG has no source difference between 64 and 256 electrodes [27].

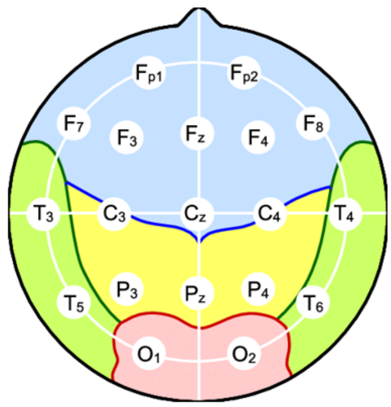


Figure 2. Conventional 10–20 EEG electrode positions for the placement of 21 electrodes [26]

2.2 Signal acquisition

An essential first step in gathering and evaluating EEG data is signal collection. Signal acquisition aims to gather comprehensible and dependable electrical impulses from the brain. The patient's features and the choice of suitable EEG machine settings determine the quality of the signals that are collected. The quality of the signal received is greatly influenced by the electrode's impedance. A high-quality electrode ensures that the signal is free of distortion and that all electrical energy at the scalp's location is transferred to the EEG equipment. The other means of transmitting the signal is required and the distance between the electrode and the scalp presents issues, particularly with the channel. Signal degradation results from this distance's electrical interference. Sweating beneath the electrode, drying of the gel electrode, and head movements by the patient can all have an impact on conduction. Uncertainty and irregularity in signals might result from these impedance aberrations. If a fatal error is made, like presuming the impedance is positive, this could have a major impact on the EEG results. It is usual practice to install an electrode in a specific region only for the purpose of assessing its impedance in order to assess the impedance and signal deterioration of the electrode. Throughout EEG treatment, routine impedance testing is frequently necessary. Reducing the impedance effect and utilizing the electrodes effectively can increase the power gain and enhance the signal quality [28].

2.3 Data processing

Data is filtered in order to eliminate noise during the cleaning process. From the sound of many nervous people to the highest level of art, this can range. This can be determined by using automated algorithms to search for patterns in numerical data that deviate from those predicted by typical brain activity. Everything that is referred to as noise is changed or eliminated. Slow data and high-frequency noise can both be eliminated from data using high-pass and low-pass filters [29].

3. FEATURE EXTRACTION METHODS

An important stage in the classification of EEG signals is feature extraction. Despite the fact that EEG signals contain literally hundreds of features that can be recovered, the feature extraction procedure for EEG signal classification can be broadly broken down into the following steps. Starting with feature selection using statistical tests like the t-test to distinguish between discriminative and non-discriminative features, and ending with dimensionality reduction using PCA. A feature selection stage particular to the classifier usually follows, albeit the feature extraction process is frequently influenced by the classification algorithm [30].

3.1 Time domain features

When separating epileptic seizure signals from regular activity, the aforementioned time domain properties can be helpful. A recent study using scalp recorded data revealed a continuous increase in line length (LL), magnitude squared area (MSA), and zero crossings of epileptic signals in comparison to normal. However, because seizure signals frequently resemble normal activity in their shape, it is reported to be challenging to distinguish them from other abnormal signals using time domain metrics. As a result, waveform analysis with wavelets has been considered for feature extraction from EEG signals [31]. A straightforward method for analyzing variations in the EEG signal's slope is zero crossing. The number of times the signal crosses the X-axis is counted to determine it. When there are positive and negative components to a signal, zero crossing can show how the waveform fluctuates between these two phases. MSA analysis was employed for this work to evaluate differences between neighboring samples. As explained in details in reference [32], signals that have low frequency with a high amplitude lead to larger MSA value and vice versa. The MSA of related EEG signal which is represented by $x(n)$ is given through Eq. (1).

$$MSA = \sum x^2(n) - x^2(n-1) \tag{1}$$

To analyzing the EEG waveform, line length was utilized to determine its movement along required x-axis. As a result, line length L for a given signal is determined by Eq. (2).

$$L = \sum |x(n) - x(n-1)| \tag{2}$$

where, the symbol |…| indicates the signal's modulus. Seizures are characterized by low frequency, high amplitude impulses that have longer line lengths than normal activity [33]. Time domain features use the time scale to characterize the properties of the EEG signal. When dealing with discrete signals, the time domain features take into account the signal value at various instants, such as $\{x(n), n = 0, \pm 1, \pm 2, \pm 3, \ldots \pm N\}$. The amplitude fluctuations of the signal above and below the baseline are used to compute the time domain characteristics. Several time domain properties, such as LL, MSA, and zero crossing, are calculated for the study of EEG signals [34].

3.2 Frequency domain features

3.2.1 Band power

It's the feature extraction technique that's most frequently employed. The power spectral density estimate method developed by Welch is used to calculate the power in each band. Once the power spectral density estimate of each frequency in a band has been added together, the power in each band is considered a feature. (Alpha: 8–13 Hz, Beta: 14–30 Hz, Gamma: >30 Hz, Theta: 4–7 Hz) [35]. Figure 3 shows the Brain wave samples with dominant frequencies belonging to different bands as more clarified in reference [36].

3.2.2 Relative band power

Its definition is the band power divided by the overall band power. When examining variations in EEG activity across several frequency bands, this function comes in handy. Generally speaking, changes in the relative band power within a certain frequency band can be used to quantify changes in activity within that band [37].

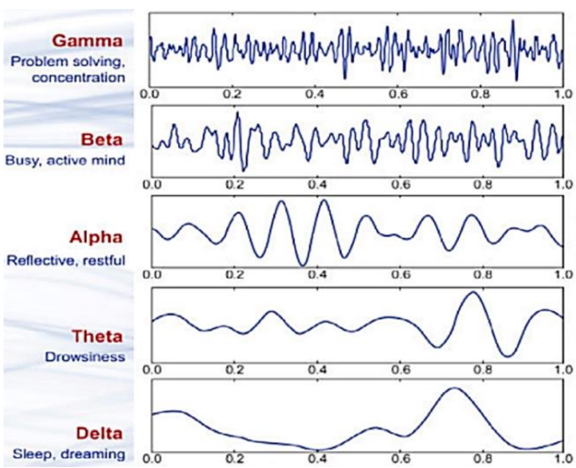


Figure 3. Brain wave samples with dominant frequencies belonging to different bands [36]

4. PROPOSED METHODOLOGY

As for the data, there are 5 types of EEG data, the size of each type is 1500×15. The transformation process was applied for this data using MATLAB to be compatible with program features. Figure 4 shows the characteristics of different data band frequencies. Dataset has been collected and transformed to the required MATLAB form [38].

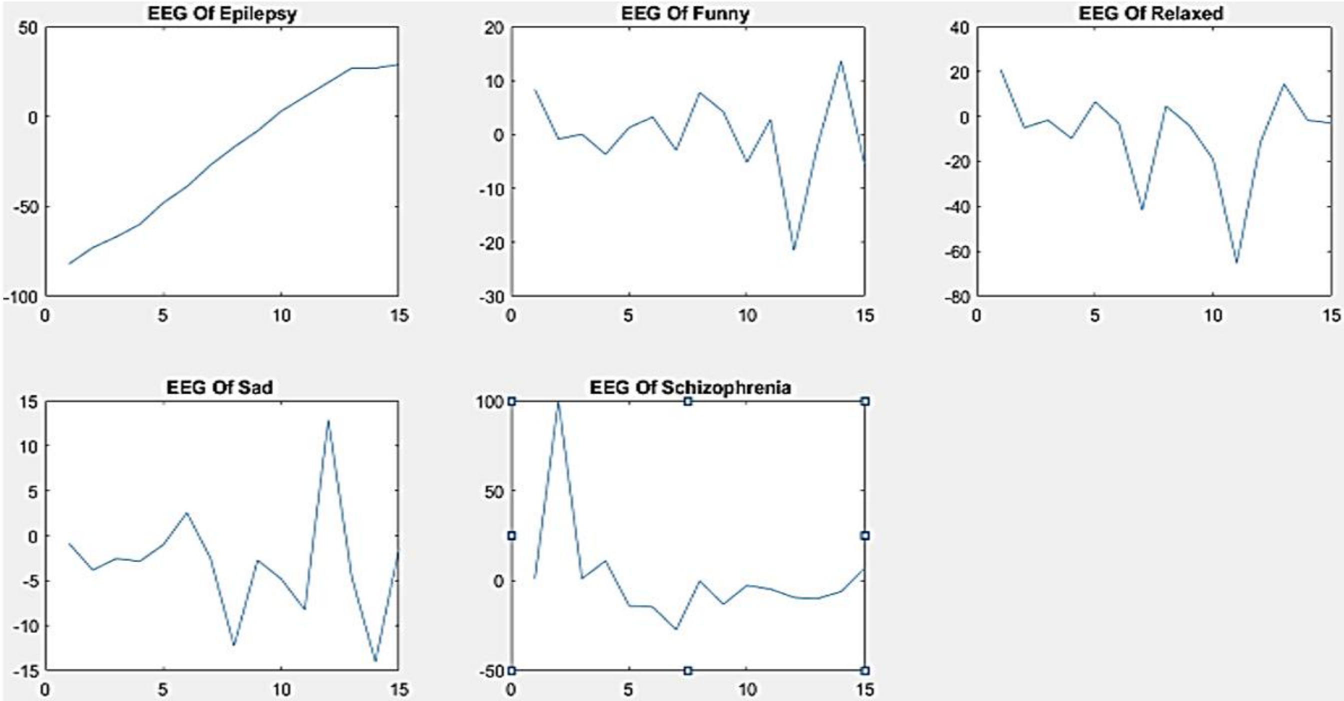


Figure 4. Emotion characteristics of the selected dataset

The flow chart that illustrates the sequential procedures involved in processing and interpreting EEG data, first create DNN with feature extraction for EEG signal categorization. The proposed flow chart that demonstrates this procedure is shown below:

1. Start
2. Pre-process EEG Signals: (removing noise, filter signals and segment into epochs).
3. Feature Extraction.
4. Feature Selection (utilization of correlation analysis technique).
5. Data Normalization (normalize features to have zero mean and unit variance).
6. Building DNN Architecture (define input, hidden and output layers' numbers).
7. Training the DNN (initialize network weights, calculate loss function, compute gradients and update weights).
8. Model Evaluation (validate on separate test dataset by measuring accuracy, time.
9. End

The proposed model flow chart is given by these steps based on EEG signal dataset:

1. Clean up the raw EEG data by segmenting into manageable epochs, applying filters, and reducing noise.
2. Extract pertinent features from pre-processed EEG signals using feature extraction. To extract various parts of signal characteristics, these features could be extracted in the time-domain, frequency-domain, or time-frequency domain.
3. To lower dimensionality and boost computational effectiveness, choose the features that are most informative.
4. To guarantee consistent scaling and promote training stability, normalize the derived characteristics.
5. Create an input layer, hidden layers (that could include recurrent or convolutional layers to capture temporal dependencies), and an output layer that is appropriate for the classification task.
6. Make use of the normalized feature vectors to train the DNN. Weights must be initialized, gradients must be computed by forward and backward propagation, and the loss function must be minimized by optimizing the model's parameters.
7. To determine the accuracy and dependability of the trained model's classification, analyse its performance using the relevant metrics on a different test dataset.

The above steps of flow chart illustrate the sequential processes from data preparation to model evaluation in a typical pipeline for processing EEG signals using a DNN with feature extraction. Variations can be implemented according to particular needs and the type of EEG signal categorization task. Defining input layers A based on input feature vector x as given in Eq. (3).

$$A^{[0]} = X \tag{3}$$

While hidden layers will be specified regarding the weight matrix (W) and bias vector (B) for a specific layer L. The activation/pre-activation output is representing using (G) and (Z) symbols respectively, while L represent the specific related layer number as in Eqs. (4) and (5).

$$Z^{[L]} = W^{[L]} A^{[L-1]} + B^{[L]} \tag{4}$$

$$A^{[L]} = G^{[L]} Z^{[L]} \tag{5}$$

Based on activation function and total layers' number N, the predicated output value will be represented as Y in Eq. (6).

$$Y = G^{[N]} Z^{[N]} \tag{6}$$

In addition to the predicated output, the loss function quantifies the total differences between predicted and total required output which is given by Eq. (7). This value was evaluated as an accuracy value for the overall prediction system. The loss function depends on (m) which is represented the number of samples and on the predicted output regarding a specific symbol i.

$$\mathrm{L_F} = -\frac{1}{\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}}(\mathrm{Y^i}\log(\mathrm{Y^i}) + (1-\mathrm{Y^i})\log(1-\mathrm{Y^i})) \tag{7}$$

Based on the model parameters, the gradient calculations of loss function enabling updating of related parameters as given for updated weight W in Eq. (8) and (9).

$$W^L = W^L - \alpha\,\frac{\partial L}{\partial W^L} \tag{8}$$

$$B^L = B^L - \alpha\,\frac{\partial L}{\partial B^L} \tag{9}$$

where, α is the learning rate and the gradient of L_F was respect to W and B that represented as $\frac{\partial L}{\partial W^L}$ and $\frac{\partial L}{\partial B^L}$ respectively. L_2 regulation technique was applied to prevent data overfitting in addition to enhance working model as given Eq. (10).

$$\mathrm{L_2} = \mathrm{L} + \frac{\lambda}{2\mathrm{m}}\sum_{\mathrm{I}=1}^{\mathrm{L}}\|\mathrm{W^L}\|_{\mathrm{F}}^2 \tag{10}$$

where, regulation parameter is λ, while the Frobenius normalization of the weight matrix is $\|W^L\|_F^2$.

Analyzing the correlation between features and the target variable can help in selecting features that have the strongest relationship with the outcome of LSTM [10, 39]. Understanding feature-target relationships can be essential for LSTMs, especially when dealing with time-series data where temporal correlations might matter. The statistical method of correlation analysis is used to quantify and examine the direction and strength of a linear relationship between variables. Correlation analysis plays a crucial role in feature selection for LSTM networks and other machine learning models by revealing which features are highly correlated with the target variable and which may be redundant. For this technique dataset should contains the desired variable (output) as well as features (input variables). While data cleaning involves handling outliers and missing values in addition to make the data has been properly preprocessed like normalization, encoding categorical variables. The rank specific utilized method was Spearman Rank Correlation. This technique was used to determines how monotonically two variables are related. It is provided a search for characteristics that have a strong positive or negative correlation with the target variable. These qualities are frequently more educational to find also the relations by examining features that have strong relationships with one another not only with classes. Excessive relation between features or even between features and classes may be a sign of feature redundancy and require dimensionality reduction. After choosing features using correlation analysis, it will help LSTM to make sure the features that have been selected will either maintain or enhance the model's performance. The process of selecting features is iterative. Based on the results of further investigation and model performance, it might need to go back and adjust LSTM initial decisions.

There were 100 units to record temporal patterns. While the Tanh was the LSTM proposed activation function, which regulates output values. While in layer 2 there were 50 units that was considering as extra LSTM layer for enhanced sequence learning. In addition, the Tanh LSTM was also the activation for this output layer. Avoiding that the model was not skewed by characteristics with higher magnitudes, the EEG data was standardized to specific range depending on values as in the range of [0, 1] using Min-Max scaling. Each EEG sample was split into windows of 1500 time steps using the sliding window technique. The sliding window method aided in capturing temporal changes between various EEG data segments. Correlation analysis was used to enhance feature quality and prevent multi-collinearity. Highly linked features were identified and eliminated using a correlation threshold of 0.9. This reduced the possibility of redundant information biasing the model. Cross-correlation was also used to identify correlations across various EEG channels over time in EEG time-series data. To find strong temporal connections between the features, a threshold of 0.7 is used.

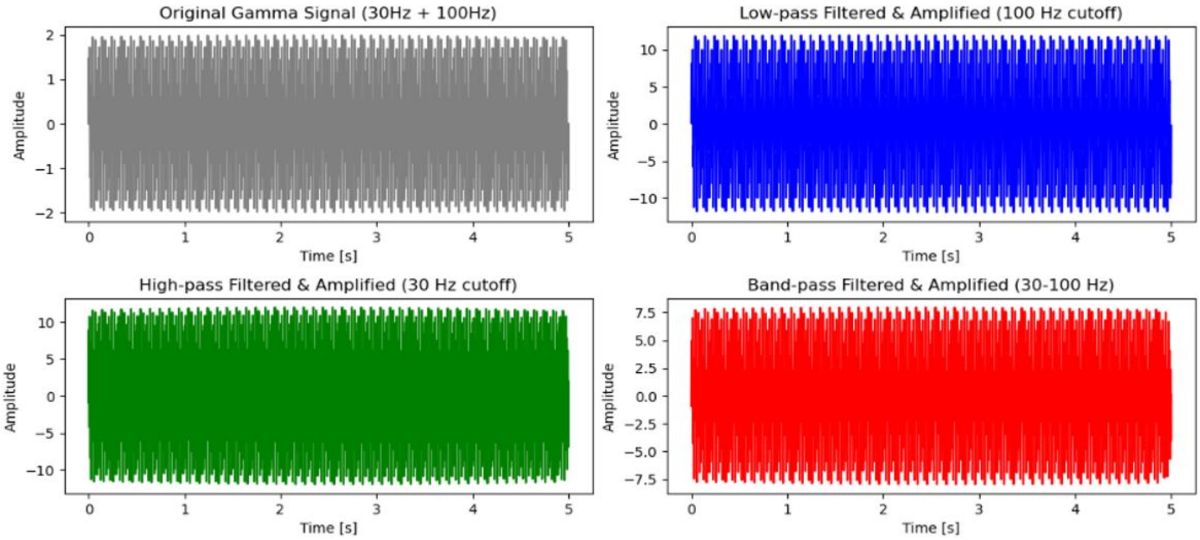
The main parameters of proposed algorithm are as follows:

1. The first layer has 100 units and the second layer has 50 units.
2. Tanh for both layers, which regulates the output of the LSTM units.
3. The EEG data was standardized using Min-Max scaling to the range [0, 1].
4. The EEG samples have been split into 1500 time-step windows for capturing temporal patterns.
5. Features with a correlation above 0.9 were removed to avoid multi-collinearity, and a threshold of 0.7 was used for cross-correlation across EEG channels.
6. Total iteration were 5800 which was used for training, with 73 epochs completed.

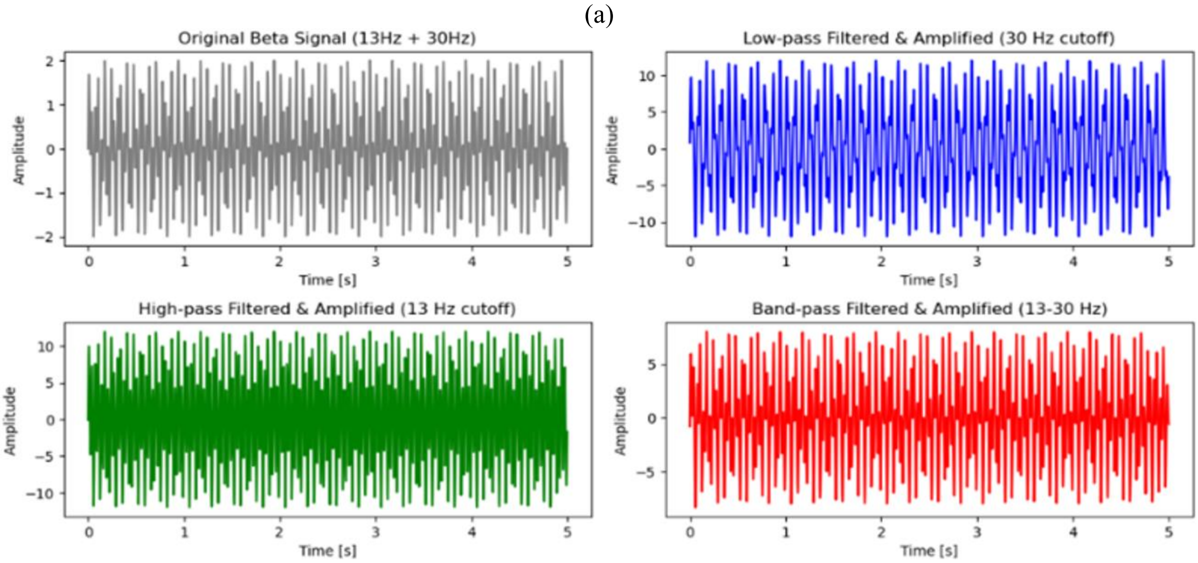
5. RESULTS

Numerous factors, including brain activity, outside inputs, and even noise from technological devices affect the complex and wide-ranging frequencies found in EEG signals. The main goal of applying filters to EEG signals was to eliminate undesirable signals while enhancing particular frequency bands that were relevant for different analysis. Filtering made it easier to analyze the data and increased the signal-to-noise ratio (SNR). Noise was reduced by using filters, particularly at particular frequencies that don't accurately reflect brain activity. For instance, low-pass filtering assisted in eliminating high-frequency noise such as muscle artifacts, whereas high-pass filtering can eliminate low-frequency drifts. Applying the right filters meaningful data has been extracted from the raw EEG signals to monitor brain activity effectively and study the relationship between brain waves and cognitive or neurological states. Before applying dataset to MATLAB, it was filtered and then amplified using Python for removing noise. Digital high-pass filter applied using tools library SciPy in Python which is built-in filter functions. Digital filtering was applied to the raw EEG dataset as part of the preprocessing stage before any learning models were deployed. Starting with Gamma which has a range between 30

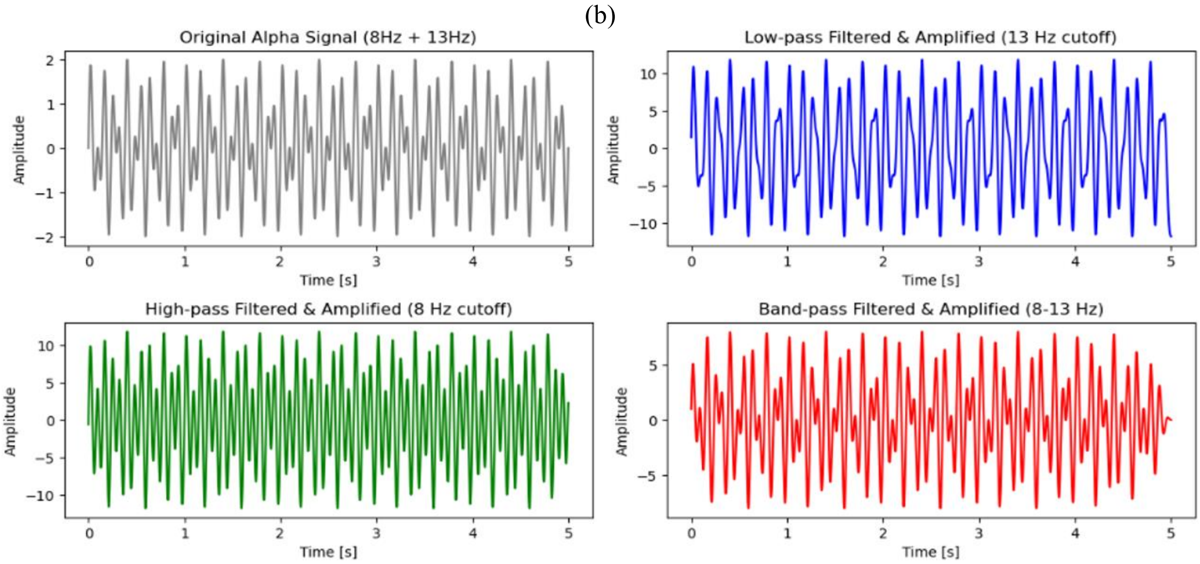
HZ and 100 HZ by applying Low, High and Band-Pass Filter with 100 HZ, 30 HZ and 30-100 HZ respectively with an amplification factor of 8. Figure 5 displayed the combined 30 Hz and 100 Hz sine waves for Gamma and other related signals. While for Beta which has a range between 13 HZ and 30 HZ, it was defined by applying Low, High and Band-Pass Filter with 30 HZ, 13 HZ and 13-30 HZ respectively with an amplification factor of 8. Figure 6 displayed the combined 30 Hz and 100 Hz sine waves. The same process was done for Alpha in range between 8-13Hz, Theta in range between 4-8 Hz and the Delta in range between 0.5-4 Hz.



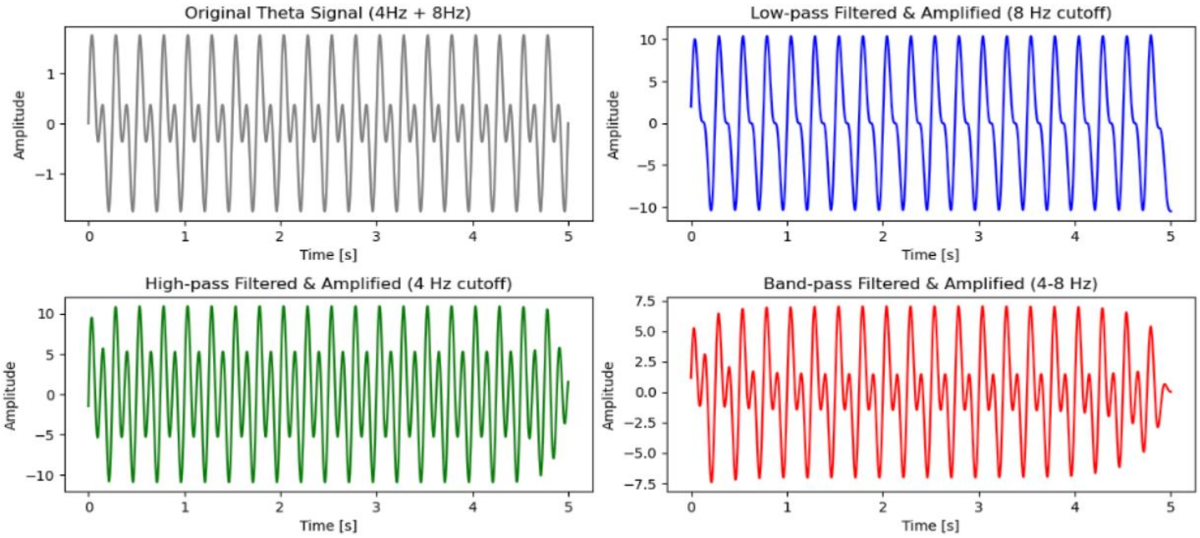
(a)



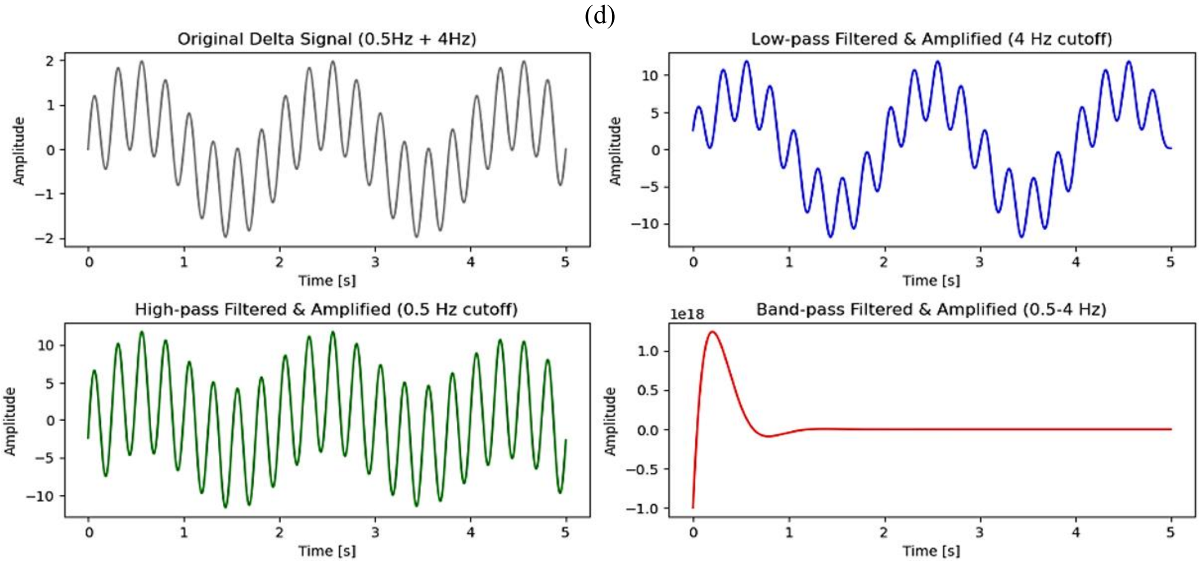
(b)



(c)



(d)



(e)

Figure 5. EEG related waves by applying Low, High and Band-pass Filters output for (a) Gamma, (b) Beta, (c) Alpha, (d) Theta and (e) Delta

MATLAB 2022b was utilized to implement the code. The data was a real-world EEG signal with five layers and an input size of 1500×15, achieving a 97.99% accuracy performance. It took 7 minutes and 58 seconds to complete the execution. The learning rate was maintained at the level of 0.005, which is normally expected. Tables 1-3 record and summarize all of the results, including a picture illustrating the differences in learning and loss between trainers and tests and the Figures 6-8. The outcomes of the DNN's training for the classification of EEG signals are highly encouraging in a number of ways as determined in Table 1 and Figure 5 for epoch 2. The network trained in 14 seconds, indicating a high level of efficiency in training, maybe due to batch processing approaches, GPU acceleration, or optimal network architecture. Two epochs, or one full pass of the training dataset, comprised the training procedure. By training throughout several epochs, the model is able to refine its weights and enhance performance by repeatedly learning from the data. Out of the potential 5800 iterations, 73 iterations were finished during training. The term iterations describe the quantity of model parameter updates resulting from the gradient descent optimization procedure. In deep learning, processing a batch of data is usually part of each iteration. However, the obtained accuracy of 97.99% shows that the trained DNN model does remarkably well in terms of EEG signal classification. The effectiveness of proposed model through 14 seconds only with epoch 2 provided a quick system testing. The well identification of EEG signals through this epoch and iteration number demonstrates the compatible of LSTM and correlation technique for such a design with 97.99%. In addition, number of iterations which was 73 iterations of 5800 as total, provided that this system can handle bigger dataset or even more complicated signals and dataset as well. Time and accuracy provided demonstration of successful utilization of LSTM with correlation technique especially for real world applications such as medical diagnosis.

Table 1. Classification performances based on LSTM + correlation at epoch 2

Standard	Value
Training Time	14 sec
Epochs	2
Iterations	73/5800
Maximum Iterations	5800
Accuracy	97.99%

Table 2. Classification performances based on LSTM + Correlation at epoch 5

Standard	Value
Training Time	33 sec
Epochs	5
Iterations	306/5800
Maximum Iterations	5800
Accuracy	98.99%

Table 3. Classification performances based on LSTM + Correlation at epoch 1000

Standard	Value
Training Time	7 min 58 sec
Epochs	1000
Iterations	5800/5800
Maximum Iterations	5800
Accuracy	99.98%

Outcomes in Table 2 and Figure 7 were obtained in this study when DNN was trained for the classification of EEG signals as well for epoch 5. The training procedure took 33 seconds to complete, demonstrating effective computing performance that was probably improved by algorithmic implementations and optimized hardware. Training was place over the course of five epochs, enabling the model to learn from the dataset iteratively. The obtained ratio from 5800 feasible iterations, 306 were carried out. This ratio of 5% from 5800 iterations with a high accuracy of about 98.99, made it easier for LSTM to change related parameters for model improvement. A reliable classification was also done for epoch 5 as in epoch 2 for EEG signal classification with high performances in addition to training time with about half minute.

Remarkable progress for LSTM with correlation technique when moving forward to 1000 epoch from 5800 iterations as shown in Table 3 and Figure 7. Training time was about eight minutes which indicate rigorous training process which effected with epochs high number. The accuracy referred to 99.98% which indicates the highest performances in EEG classification among other epochs number and iterations. Overall results provided LSTM flexibility with feature selection methods and demonstrate that system could extract small patterns for input signals such as EEG signal.

Table 4. Comparison system performances based on different dataset with similar/non-similar Techniques

Reference No.	Utilized Techniques	Accuracy	Information
[5]	DNN	< 97%	SaHeart (SHt) dataset
[5]	LSTM	< 94%	SaHeart (SHt) dataset
[5]	RNN	< 93%	Wisconsin Breast Cancer dataset
[40]	CNN + ReLU	87.3%	DEAP dataset
[41]	CNN	82%	With feature selection
[42]	DNN	83.98%	1- IIa of BCI Competition IV dataset
[42]	DNN	83.98%	2- IIb of BCI Competition IV dataset
[43]	CNN + ReLU	79.3%	BCI Competition IV dataset
[43]	CNN + ReLU	85.7%	For TJU dataset
[44]	LSTM	61.08 %	Simultaneous Task EEG Workload
[44]	CNN+ LSTM	58.68%	Simultaneous Task EEG Workload
[45]	CNN + ReLU	90%	With feature selection
Proposed method	LSTM + Correlation as feature selection	97.99%	Kaggle dataset for epoch 2
Proposed method	LSTM + Correlation as feature selection	98.99%	Kaggle dataset for epoch 5
Proposed method	LSTM + Correlation as feature selection	99.98%	Kaggle dataset for epoch 1000

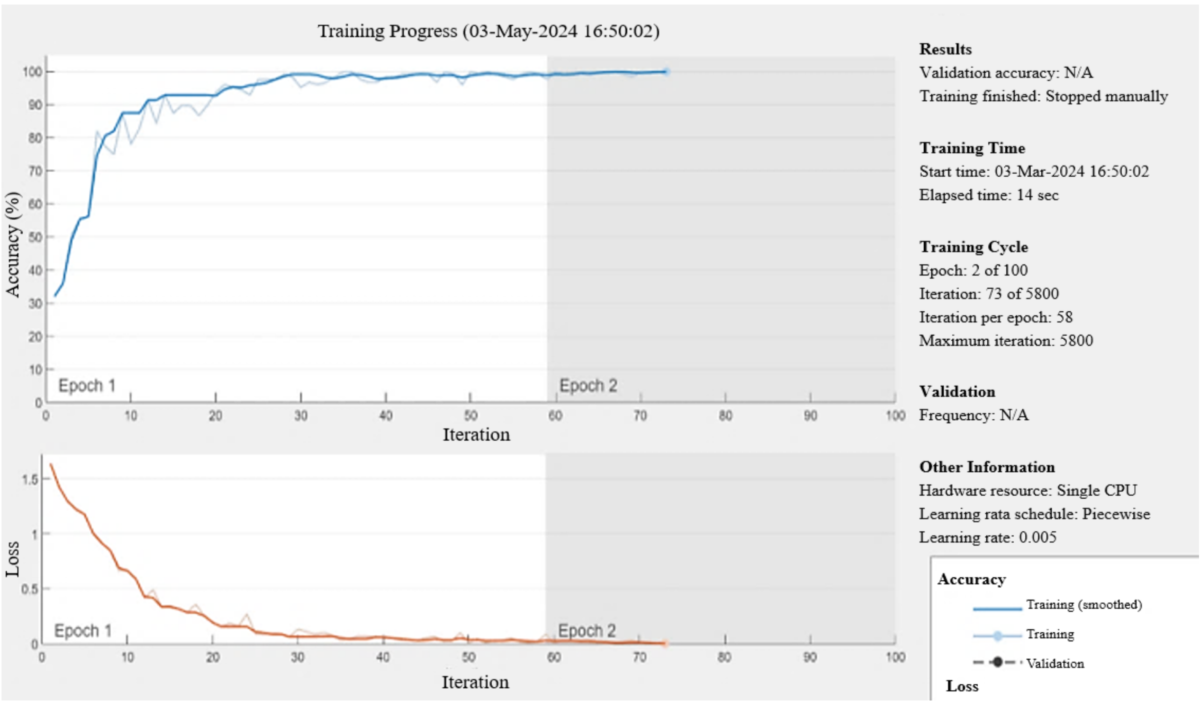


Figure 6. The DNN's performances with loss and accuracy of EEG signals classification at epochs 2

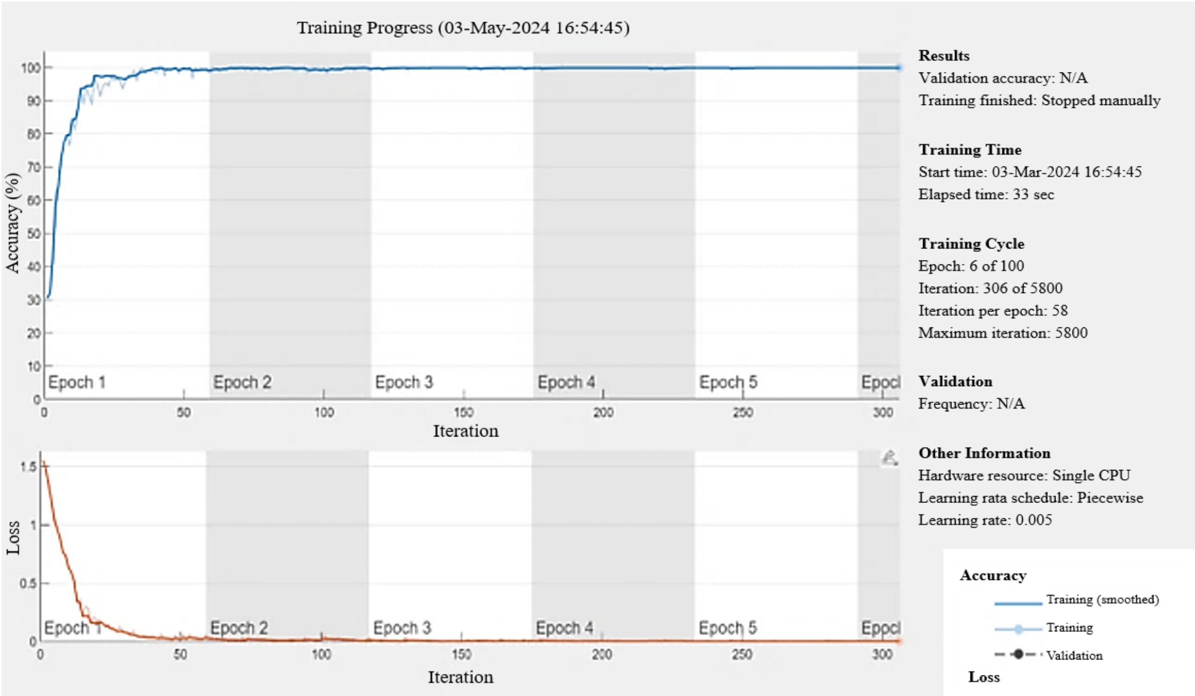


Figure 7. The DNN's performances with loss and accuracy of EEG signals classification at epochs 5

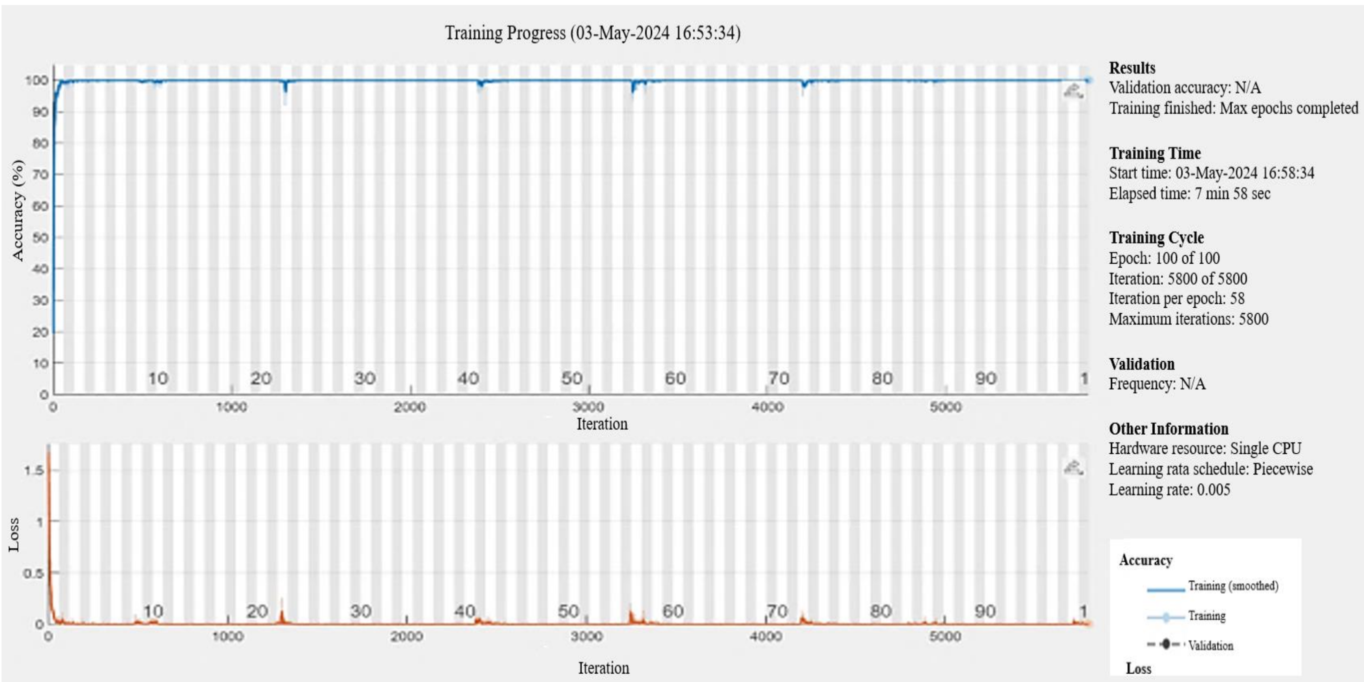


Figure 8. The DNN's performances with loss and accuracy of EEG signals classification at epochs 1000

Table 4 shows comparison between several topologies related to the same field for a similar task with the proposed model in this work. The accuracy of compared articles was in range of 83 and 97 which demonstrate the stability privileges of DNN. While combining LSTM with a feature selection technique provide a higher range between 61 and 99, that lead to the fact of LSTM need a better matching for feature techniques than others. An older version of LSTMs called RNNs had accuracy levels below 93%, highlighting its shortcomings in comparison to more sophisticated designs. While depending on how they were configured, CNNs produced a range of results. For example, CNNs using Rectified Linear Unit (ReLU) activations achieved accuracies of 87.3% to 90%, demonstrating their efficacy in tasks combining pattern recognition and spatial data. The accuracy of a CNN plus LSTM combination, however, was much lower at 58.68%, indicating difficulties in successfully combining these architectures for the given task. A notable result of feature engineering's substantial influence on model performance is the high accuracies that LSTM models improved with feature selection techniques routinely attained, ranging from 97.99% to 99.98%. Neural topologies need a fine-tuning in their parameters in addition to match the suitable case of feature selection for such an application like EEG classification. Compared to CNN + ReLU methods, LSTM provided best results in this work when combing it with correlation technique. Particularly when improved with efficient feature selection strategies, demonstrating their usefulness in jobs that necessitated long-term dependency modeling and sequential data processing.

Important performance indicators were also indicated like precision, recall, and F1-score in addition to the model's accuracy as shown in Tables 1-3 which is related to epochs values 2, 5, and 1000 respectively. Beyond basic accuracy, which can occasionally be deceptive, particularly in imbalanced datasets, these measures offer a more thorough understanding of the model capacity to classify the data. At epoch 2 the accuracy was 97.99%, precision value was 97. 40%, the recall value was 98.10%, while the F1-Score was 97.75%. The model performed comparatively well on all measuring criteria at epoch 2. The model was successfully balancing false positives and false negatives, as evidenced by the close precision and recall values at this epoch. Additionally, the F1-score was high, indicating well-maintained trade-off between recall and precision.

Also, at epoch 5 the accuracy was 98.99%, precision value was 98.50%, the recall value was 99.10%, while the F1-Score was 98.80%. The model was getting better at recognizing both the positive and negative classes, as evidenced by the recall improving by 1.00% and the precision increasing by roughly 1.10%. The better balance between recall and precision was further supported by the rise in F1-score. In addition, at epoch 1000, the accuracy was 99.98%, the precision value was 99.80%, while the recall and F1-Score were 99.90% and 99.85% respectively. The model achieved near-perfect accuracy and a significant improvement in precision and recall at Epoch 1000. With an F1-score of 99.85%, the model had significantly improved over previous epochs and was very good at categorizing the data with few false positives and false negatives. Statistical tests have been conducted to see whether the reported variations in accuracy, precision, recall and F1-score between epoch 2, epoch 5 and epoch 1000 are statistically significant in order to validate the observed performance increases. Paired t-test has been utilized to compare the accuracy values across epochs. The p-value was 0.003 between epochs 2 and 5 that showed a highly improvement. Also, the p-value was 0.0001 between epoch 5 and 1000 which is also showed significant improvement.

The accuracy obtained for this model between these epochs were statistically significant because the p-values were less than the often-used cutoff of 0.05. In addition, the non-parametric Wilcoxon signed-rank test was applied for precision, recall and F1-score to demonstrate that the outcome was not the result of chance variation. The results showed that precision between epochs 2 and 5 was 0.004, Recall between epochs 2 and 5 was 0.005, F1-Score comparison between epochs 2 and 5 was 0.003. In addition, the Precision between epoch 5 and 1000 was 0.0003, the Recall between 5 and 1000 was 0.0001 and F1-Score for 5 vs 1000 was 0.0002. The statistical significance of the improvements in precision, recall and F1-score between the epochs was confirmed by the fact that all p-values are significantly below 0.05. The results showed that when training time which is indicated to epoch increases, model performance clearly improved. The measuring metrics were all increased, indicating that the model enhanced over time at correctly classifying data with diminishing returns beyond the initial epochs. The fact that results were more gradual between Epoch 5 and 1000, although being significant between Epochs 2 and 5, was indicative of this. The statistical tests confirmed that the observed improvements were meaningful and not the result of chance variations.

EEG signals have long-term temporal relationships and are sequential by nature and these kinds of dependencies are frequently too difficult for traditional feedforward neural networks to capture. LSTM networks are perfect for simulating EEG signals since they were specifically made to handle sequential data and reduce problems like the vanishing gradient problem where the data order is crucial. Iterations with 73 were finished during training through 5800 iterations. Using correlation-based feature selection was one of the main ways the suggested model prevents overfitting. The model was shielded from learning from redundant data by removing strongly correlated characteristics, those with correlations greater than 0.9. The model was less likely to learn repetitive patterns that were not transfer well to fresh data. It was also compelled to concentrate on the most instructive signals in the EEG data by decreasing the complexity of the feature space and eliminating superfluous features. Additionally, it kept keeps the model from getting overly complicated which could result in overfitting. The model capacity to generalize to previously unseen data was improved by having fewer features, which made it less susceptible to fitting noise and irrelevant patterns. The underlying structure of LSTMs, which were proposed to capture long-term dependencies in EEG, comprises mechanisms to regulated the information flow. The LSTM has the capacity to forget unimportant information which reduced the probability of overfitting to erroneous patterns in the data by helping to concentrate on the important patterns. LSTMs are more resistant to overfitting than typical feed-forward networks, which are prone to overfitting when presented with sequential data.

Cross-validation was a crucial step in preventing overfitting during training which deduced from the outcomes. The process of cross-validation involved dividing the dataset into several subsets and then training and assessing the model on several training/testing splits. This lowered the possibility of overfitting to a single subset of the data and aided in evaluating how well the model generalizes to various subsets of the data. LSTM networks are specifically made to capture long-range relationships across time, traditional feed-forward neural networks such as DNNs are unable to model these temporal dependencies since they consider the data as separate samples. Brain activity at one time point is frequently tightly linked to activity at earlier time points in EEG readings. In order to capture the sequential structure in the data, LSTMs are particularly good at remembering and forgetting significant elements of these sequences. Although both RNNs and LSTMs are made for sequential data, LSTMs performed better than simple RNNs when dealing with long-term dependencies because they can prevent vanishing gradients during training. Learning across lengthy sequences can be affected by the vanishing gradient issue in conventional RNNs. However, LSTMs addressed this issue, resulting in more reliable and efficient learning over larger time periods in EEG data. Overfitting may result from the high correlation between several features like distinct EEG channels in EEG datasets. By eliminating unnecessary variables, correlation-based feature selection assisted the model concentrate on the most important and independent features. This enhanced the model's capacity for generalization which enable it to function better on data that has not been seen yet.

In the absence of feature selection, LSTM networks became overfit due to an abundance of redundant or irrelevant information. The model was made simpler and more targeted by eliminating strongly correlated features with a threshold value of 0.9 and keeping only the most pertinent data elements. This lowered the possibility of overfitting and enhances generalization to test data. Since CNN and CNN + ReLU models did not usually carry out feature selection in the previous research, they overfitted to particular data points. Additionally, they performed worse on sequential data since they are unable to model the temporal structure of EEG signals as well as LSTM networks. Strong convergence and good generalization are indicated by the suggested model persistent near-perfect accuracy over several training sessions and epochs due to previous reasons. The sequential modeling of LSTM and correlation-based feature selection probably improved the model ability to generalize to new or unseen data, even though overfitting is a possible with such high accuracy. The model appeared to be learning efficiently with minimal overfitting, as evidenced by the good accuracy after 1000 epochs. The model was learning strong, generalizable patterns rather than merely memorizing the data as evidenced by the steady rise in accuracy over time from 97.99% to 99.98%. The reported accuracy represented consistent performance across several data subsets rather than being the product of overfitting to a single training set.

The suggested LSTM + Correlation feature selection strategy worked very well since it reduced redundancy through feature selection while capturing temporal dependencies in EEG data. Even though the stated accuracy of 99.98% was remarkable, methods such as correlation-based feature selection and the intrinsic characteristics of LSTM networks were used to omit overfitting. The model improved generalization and lowered the chance of overfitting by removing superfluous features.

6. CONCLUSION

DNN for its first training phase across two epochs, completing 73 iterations out of a maximum of 5800. Training time for 2 epochs was short and provided a quick processing due to parameters good tuning between feature selection and LSTM technique. Few trainings iteration as well provided this result with about 97.99% and lead to the successful utilization of proposed model to classify EEG signal. Advanced epoch to 5 took about half minute with still effective training process. Moving through dataset in several times with about 306 from higher iteration total number, system was able to deep learning with about 98.99% signifying enhanced proficiency in differentiating between EEG signal patterns. Understanding linkage in EEG features was better when adding more epochs which lead to classification improvement. For 1000 epochs, the training time was increased as well to 8 minutes approximately, which made this model fine tuning with related signal and LSTM parameters. The final accuracy was 99.98%, which was over long learning time due to identify minute changes in EEG data. This system can be applied for accurate categorization such as diagnosis, As the number of training epochs increased, accuracy increased steadily, demonstrating the progressive advantages of extended training in improving the model's prediction abilities. While in iterations advanced from the first stages to the final stage, the model's capacity for adaptation was strengthened over time from seconds to minutes. These comprehensive results highlight the importance of training depth in developing DNN-based EEG signal classification systems by demonstrating performance increases attained through varied training durations and epoch from 2 to 1000. Pre-processing with feature selection and LSTM well tuning parameters are the three main components of the analytic process in this work. Furthermore, LSTM technique was created in this work to carry out the classification process, which was simple to implemented by MATLAB R2022b. When DNN is used for the training and testing of the extracted features, the suggested methodology obtains an average classification accuracy of 97.99%, 98.99%, and 99.98%, respectively with second to minutes training time.

This model has a wide range of possible clinical uses. EEG is frequently utilized in medical contexts for the monitoring and diagnosis of neurological illnesses, including sleep disorders and brain-computer interface (BCI) applications. Based on the results, the suggested model may be applied to real-time monitoring devices to identify abnormalities in brain activity, giving medical professionals important new information. For example, by more accurately detecting aberrant EEG patterns than existing systems, this model may help in the early detection of neurological illnesses. Additionally, the model's scalability is shown by its capacity to handle big datasets as seen by its performance with up to 5800 iterations, which makes it appropriate for implementation in major healthcare systems where processing and analysis of EEG data from several patients is required. The short training time with 14 seconds for epoch 2 yielded outstanding performance that suggests that this model may be used to real-world clinical settings that need low latency and efficiency, including real-time patient monitoring in emergency rooms or intensive care units. The use of this concept may also be expanded to brain-computer interfaces and neuro-prosthetics, where external devices are controlled by EEG signals. Because of the model's exceptional ability to identify EEG signals, people with impairments may be able to use assistive devices and prostheses more precisely that would enhance their quality of life. A number of restrictions must be noted for the proposed model which were the model performance was dependent on a particular taken dataset and although the outcomes were promising, it was yet unclear whether the model could be applied to other datasets. To evaluate the model's resilience across various EEG signal types, more validation on larger and diverse datasets including actual patient data is required because the EEG signals in the Kaggle dataset could not be the same as those seen in other experimental situations. Second, even if the accuracy was good, training it takes a lot of computer power, especially at larger epochs like 1000 epochs. This makes it less feasible to use the model in settings with limited resources. Even while the training duration of about 8 minutes at epoch 1000 is quick for complicated models, in some situations it can be too long for real-time applications. For clinical application, it would be essential to optimize the model for quicker inference and less extensive training. The dependence on the correlation feature selection method, which might not be generally applicable to all EEG signal types was another drawback. The particulars of the dataset and the feature selection technique could have a significant impact on the model's performance. Alternative feature engineering approaches and their effects on model performance should be investigated in future research. The interpretability of AI models is essential in therapeutic contexts. Future research might concentrate on increasing the models decision-making transparency so that physicians can comprehend how the model makes its predictions. In order to improve trust and usability in practical applications, this may entail incorporating explainable AI techniques to offer insights into the characteristics that most influence the model predictions. When aberrant brain activity is identified, integrating this model into real-time EEG monitoring devices may enable prompt clinical intervention. This could entail combining the model with implanted brain sensors or wearable EEG equipment to provide a continuous and noninvasive way to track brain health. The paper also made another contribution by showing how effective the suggested model was in terms of execution and training time. With epoch 2 taking only 14 seconds, epoch 5 taking 33 seconds and epoch 1000 taking 7 minutes 58 seconds, the model demonstrated its ability to train quickly. These outcomes demonstrated the model computational effectiveness, which qualifies it for real-time applications where speed is essential. The model demonstrated its potential for useful, large-scale data processing by maintaining good performance even as the number of epochs and iterations increased. This comparative analysis offered important insights for future signal processing research by highlighting the significance of choosing suitable feature extraction techniques to improve model performance in addition to proving the efficacy of LSTM for EEG classification tasks. This study also demonstrated the adaptability and strength of LSTM networks, especially when combined with feature selection methods. As evidenced by the steady increase in accuracy over several epochs, the model capacity to fine-tune its parameters demonstrated how adaptable LSTM was while processing intricate sequential data. When working with data like EEG signals, which frequently contain non-linear patterns and long-term dependencies, this flexibility was very crucial. This research had significant practical ramifications, the model classification accuracy pointed to its potential for usage in real-time applications, like ongoing EEG signal monitoring for the early identification of neurological illnesses. Additionally, the system used in clinical situations, where prompt choices are crucial, it made possible by capacity to analyze huge datasets efficiently and rapid training rate. This work demonstrated the possibility for additional advancement in the use of deep learning models to solve practical issues, especially in the medical field. Future research might examine the model scalability to even bigger and more varied datasets and look into how it can be integrated with other signal processing methods to improve its resilience and cross-domain applicability.

REFERENCES

[1] Panteliadis, C.P. (2021). Historical overview of electroencephalography: From antiquity to the beginning of the 21st Century. Journal of Brain and Neurological Disorders, 3(1): 1-10. https://doi.org/10.5281/zenodo.5359323

[2] Arora, N., Mishra, B. (2021). Origins of ECG and evolution of automated DSP techniques: A review. IEEE Access, 9: 140853-140880. https://doi.org/10.1109/ACCESS.2021.3119630

[3] Galván, E., Mooney, P. (2021). Neuroevolution in deep neural networks: Current trends and future challenges. IEEE Transactions on Artificial Intelligence, 2(6): 476-493. https://doi.org/10.1109/TAI.2021.3067574

[4] Adate, A., Tripathy, B.K., Arya, D., Shaha, A. (2020). Impact of deep neural learning on artificial intelligence research. Deep Learning Research and Applications, 7: 69-84. https://doi.org/10.1515/9783110670905

[5] Bikku, T. (2020). Multi-layered deep learning perceptron approach for health risk prediction. Journal of Big Data, 7(1): 50. https://doi.org/10.1186/s40537-020-00316-7

[6] Ding, D.D., Ma, Z., Chen, D., Chen, Q.S., Liu, Z., Zhu, F.Q. (2021). Advances in video compression system using deep neural network: A review and case studies. Proceedings of the IEEE, 109(9): 1494-1520. https://doi.org/10.1109/JPROC.2021.3059994

[7] Khanafer, M., Shirmohammadi, S. (2020). Applied AI in instrumentation and measurement: The deep learning revolution. IEEE Instrumentation & Measurement Magazine, 23(6): 10-17. https://doi.org/10.1109/MIM.2020.9200875

[8] Wang, X., Zhao, Y., Pourpanah, F. (2020). Recent advances in deep learning. International Journal of Machine Learning and Cybernetics, 11: 747-750. https://doi.org/10.1007/s13042-020-01096-5

[9] Xiao, G., Li, J., Chen, Y., Li, K. (2020). MalFCS: An effective malware classification framework with automated feature extraction based on deep convolutional neural networks. Journal of Parallel and Distributed Computing, 141: 49-58. https://doi.org/10.1016/j.jpdc.2020.03.012

[10] Mohammed, S.J., Mohammed, M.S. (2022). COVID-19 risk factors specification using Decision Tree based on the degree of redundancy between features. In 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT), Bangalore, India, pp. 1-11. https://doi.org/10.1109/GCAT55367.2022.9971950

[11] Saeidi, M., Karwowski, W., Farahani, F.V., Fiok, K., Taiar, R., Hancock, P.A., Al-Juaid, A. (2021). Neural decoding of EEG signals with machine learning: A systematic review. Brain Sciences, 11(11): 1525. https://doi.org/10.3390/brainsci11111525

[12] Hosseini, M.P., Hosseini, A., Ahi, K. (2020). A review on machine learning for EEG signal processing in bioengineering. IEEE Reviews in Biomedical Engineering, 14: 204-218. https://doi.org/10.1109/RBME.2020.2969915

[13] Tosun, M., Çetin, O. (2022). A new phase-based feature extraction method for four-class motor imagery classification. Signal, Image and Video Processing, 16: 1-8. https://doi.org/10.1007/s11760-021-02035-9

[14] Merlin Praveena, D., Angelin Sarah, D., Thomas George, S. (2022). Deep learning techniques for EEG signal applications—A review. IETE journal of Research, 68(4): 3030-3037. https://doi.org/10.1080/03772063.2020.1749143

[15] Guerrero, M.C., Parada, J.S., Espitia, H.E. (2021). Principal components analysis of EEG signals for epileptic patient identification. Computation, 9(12): 133. https://doi.org/10.3390/computation9120133

[16] Guo, J., Xu, T., Xie, L., Liu, Z. (2024). Exploring an intelligent classification model for the recognition of automobile sounds based on EEG physiological signals. Mathematics, 12(9): 1297. https://doi.org/10.3390/math12091297

[17] Al Fahoum, A. (2023). Early detection of neurological abnormalities using a combined phase space reconstruction and deep learning approach. Intelligence-Based Medicine, 8: 100123.

https://doi.org/10.1016/j.ibmed.2023.100123
[18] Qin, S., Ji, Z. (2004). Multi-resolution time-frequency analysis for detection of rhythms of EEG signals. 3rd IEEE Signal Processing Education Workshop. 2004 IEEE 11th Digital Signal Processing Workshop, 2004. Taos Ski Valley, USA, pp. 338-341. https://doi.org/10.1109/DSPWS.2004.1437971
[19] Cataldo, A., Criscuolo, S., De Benedetto, E., Masciullo, A., Pesola, M., Picone, J., Schiavoni, R. (2024). EEG complexity-based algorithm using multiscale fuzzy entropy: Towards a detection of Alzheimer's disease. Measurement, 225: 114040. https://doi.org/10.1016/j.measurement.2023.114040
[20] Ke, H., Chen, S., Zhang, H., Tang, Y., Liu, Y., Chen, D., Li, X. (2017). A shallow-dense network approach to synchronization pattern classification of multivariate epileptic EEG. In Chinese Intelligent Systems Conference, Springer, Singapore, pp. 553-563. https://doi.org/10.1007/978-981-10-6496-8_51
[21] Koenig, T., Hubl, D., Mueller, T.J. (2002). Decomposing the EEG in time, space and frequency: A formal model, existing methods, and new proposals. International Congress Series, 1232: 317-321. https://doi.org/10.1016/S0531-5131(01)00724-5
[22] Miinalainen, T., Rezaei, A., Us, D., Nüßing, A., Engwer, C., Wolters, C.H., Pursiainen, S. (2019). A realistic, accurate and fast source modeling approach for the EEG forward problem. NeuroImage, 184: 56-67. https://doi.org/10.1016/j.neuroimage.2018.08.054
[23] Kalus, S., Sämann, P., Czisch, M., Fahrmeir, L. (2013). fMRI activation detection with EEG priors. Technical Report Number 146, Department of Statistics, University of Munich. https://doi.org/10.5282/ubm/epub.15725
[24] Lei, X., Valdes-Sosa, P.A., Yao, D. (2012). EEG/fMRI fusion based on independent component analysis: Integration of data-driven and model-driven methods. Journal of Integrative Neuroscience, 11(03): 313-337. https://doi.org/10.1142/S0219635212500203
[25] Asadzadeh, S., Rezaii, T.Y., Beheshti, S., Delpak, A., Meshgini, S. (2020). A systematic review of EEG source localization techniques and their applications on diagnosis of brain abnormalities. Journal of Neuroscience Methods, 339: 108740. https://doi.org/10.1016/j.jneumeth.2020.108740
[26] Bos, D.O. (2006). EEG-based emotion recognition. The Influence of Visual and Auditory Stimuli, 56(3): 1-17.
[27] Ang, K.K., Chua, K.S.G., Phua, K.S., Wang, C., Chin, Z. Y., Kuah, C.W.K., Low, W., Guan, C. (2015). A randomized controlled trial of EEG-based motor imagery brain-computer interface robotic rehabilitation for stroke. Clinical EEG and Neuroscience, 46(4): 310-320. https://doi.org/10.1177/1550059414522229
[28] Hinrichs, H., Scholz, M., Baum, A.K., Kam, J.W., Knight, R.T., Heinze, H.J. (2020). Comparison between a wireless dry electrode EEG system with a conventional wired wet electrode EEG system for clinical applications. Scientific Reports, 10(1): 5218. https://doi.org/10.1038/s41598-020-62154-0
[29] Ghosh, R., Phadikar, S., Deb, N., Sinha, N., Das, P., Ghaderpour, E. (2023). Automatic eyeblink and muscular artifact detection and removal from EEG signals using k-nearest neighbor classifier and long short-term memory networks. IEEE Sensors Journal, 23(5): 5422-5436. https://doi.org/10.1109/JSEN.2023.3237383
[30] Wang, J., Wang, M. (2021). Review of the emotional feature extraction and classification using EEG signals. Cognitive Robotics, 1: 29-40. https://doi.org/10.1016/j.cogr.2021.04.001
[31] Al-Qerem, A., Kharbat, F., Nashwan, S., Ashraf, S., Blaou, K. (2020). General model for best feature extraction of EEG using discrete wavelet transform wavelet family and differential evolution. International Journal of Distributed Sensor Networks, 16(3): 1550147720911009. https://doi.org/10.1177/1550147720911009
[32] Wen, T., Narita, F., Kurita, H., Jia, Y., Shi, Y. (2023). Quantification of damage expansion influence on frequency response function of plate for structural health monitoring with integral differential method. Composites Science and Technology, 244: 110298.
[33] Adamatzky, A. (2022). Language of fungi derived from their electrical spiking activity. Royal Society Open Science, 9(4): 211926. https://doi.org/10.1098/rsos.211926
[34] Anuragi, A., Sisodia, D.S., Pachori, R.B. (2021). Automated FBSE-EWT based learning framework for detection of epileptic seizures using time-segmented EEG signals. Computers in Biology and Medicine, 136: 104708. https://doi.org/10.1016/j.compbiomed.2021.104708
[35] Anthony, R.E., Ringler, A.T., Wilson, D.C., Bahavar, M., Koper, K.D. (2020). How processing methodologies can distort and bias power spectral density estimates of seismic background noise. Seismological Research Letters, 91(3): 1694-1706. https://doi.org/10.1785/0220190212
[36] Kumar, P., Abubakar, A.A., Sazili, A.Q., Kaka, U., Goh, Y.M. (2022). Application of electroencephalography in preslaughter management: A review. Animals, 12(20): 2857. https://doi.org/10.3390/ani12202857
[37] Donoghue, T., Haller, M., Peterson, E.J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, J.D., Knight, R.T., Shestyuk, A., Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. Nature Neuroscience, 23(12): 1655-1665. https://doi.org/10.1038/s41593-020-00744-x
[38] EEG-Alcohol. (2017). https://www.kaggle.com/datasets/nnair25/Alcoholics?resource=download.
[39] Mohammed, S.J.M., Ahmed, A.A., Ahmad, A.A., Mohammed, M.S. (2020). Anemia prediction based on rule classification. In 2020 13th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, United Kingdom, pp. 427-431. https://doi.org/10.1109/DeSE51703.2020.9450234
[40] Qiao, R., Qing, C., Zhang, T., Xing, X., Xu, X. (2017). A novel deep-learning based framework for multi-subject emotion recognition. In 2017 4th International Conference on Information, Cybernetics and Computational Social Systems (ICCSS), Dalian, China, pp. 181-185. https://doi.org/10.1109/ICCSS.2017.8091408
[41] Morabito, F.C., Campolo, M., Ieracitano, C., Ebadi, J.M., Bonanno, L., Bramanti, A., Desalvo, S., Mammone, N., Bramanti, P. (2016). Deep convolutional neural networks for classification of mild cognitive impaired and Alzheimer's disease patients from scalp EEG recordings.

In 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI), Bologna, Italy, pp. 1-6. https://doi.org/10.1109/RTSI.2016.7740576
[42] Zhao, H., Zheng, Q., Ma, K., Li, H., Zheng, Y. (2020). Deep representation-based domain adaptation for nonstationary EEG classification. IEEE Transactions on Neural Networks and Learning Systems, 32(2): 535-545. https://doi.org/10.1109/TNNLS.2020.3010780
[43] Sun, B., Zhao, X., Zhang, H., Bai, R.F., Li, T. (2020). EEG motor imagery classification with sparse spectrotemporal decomposition and deep learning. IEEE Transactions on Automation Science and Engineering, 18(2): 541-551. https://doi.org/10.1109/TASE.2020.3021456
[44] Pandey, V., Choudhary, D.K., Verma, V., Sharma, G., Singh, R., Chandra, S. (2020). Mental workload estimation using EEG. In 2020 Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Bangalore, India, pp. 83-86. https://doi.org/10.1109/ICRCICN50933.2020.9296150
[45] Jiao, Z.C., Gao, X.B., Wang, Y., Li, J., Xu, H.J. (2018). Deep convolutional neural networks for mental load classification based on EEG data. Pattern Recognition, 76: 582-595. https://doi.org/10.1016/j.patcog.2017.12.002

NOMENCLATURE

A	input layers
x	input feature vector
G	activation output
Z	pre-activation output
N	layers' number
Y	predicated output value
m	number of samples

Greek symbols

α	the learning rate
λ	regulation parameter