mathematics

MDPI

Article

Automated EEG Pathology Detection Based on Significant Feature Extraction and Selection

Yunning Zhong [1,†], Hongyu Wei [2,†], Lifei Chen [2,*] and Tao Wu [1,*]

[1] School of Mathematics and Statistics, Fujian Normal University, Fuzhou 350117, China
[2] College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China
* Correspondence: clfei@fjnu.edu.cn (L.C.); qbx20210075@yjs.fjnu.edu.cn (T.W.)
† These authors contributed equally to this work.

Abstract: Neurological diseases are a significant health threat, often presenting through abnormalities in electroencephalogram (EEG) signals during seizures. In recent years, machine learning (ML) technologies have been explored as a means of automated EEG pathology diagnosis. However, existing ML-based EEG binary classification methods largely focus on extracting EEG-related features, which may lead to poor performance in classifying EEG signals by overlooking potentially redundant information. In this paper, we propose a novel Kruskal–Wallis (KW) test-based framework for EEG pathology detection. Our framework first divides EEG data into frequency sub-bands using wavelet packet decomposition and then extracts statistical characteristics from each selected coefficient. Next, the piecewise aggregation approximation technique is used to obtain the aggregated feature vectors, followed by the KW statistical test methodology to select significant features. Finally, three ensemble learning classifiers, random forest, categorical boosting (CatBoost), and light gradient boosting machine, are used to classify the extracted significant features into normal or abnormal classes. Our proposed framework achieves an accuracy of 89.13%, F1-score of 87.60%, and G-mean of 88.60%, respectively, outperforming other competing techniques on the same dataset, which shows the great promise in EEG pathology detection.

Keywords: pathology detection; significant features; Kruskal–Wallis test; wavelet packet decomposition; electroencephalography

MSC: 68T01; 92C55

Citation: Zhong, Y.; Wei, H.; Chen, L.; Wu, T. Automated EEG Pathology Detection Based on Significant Feature Extraction and Selection. *Mathematics* **2023**, *11*, 1619. https://doi.org/10.3390/math11071619

Academic Editors: Guohun Zhu, Lemai Nguyen and Xujuan Zhou

Received: 7 March 2023
Revised: 24 March 2023
Accepted: 25 March 2023
Published: 27 March 2023

1. Introduction

The brain is a crucial organ responsible for receiving external stimuli, generating sensations, forming consciousness, and providing instructions. Therefore, research on the brain has been a subject of immense interest. In recent decades, electroencephalogram (EEG) has become a useful auxiliary tool for brain research. It allows researchers to record and monitor spontaneous brain activity. Unlike other existing biomedical imaging techniques, such as magnetic resonance imaging [1], EEG offers numerous advantages, including non-invasiveness, high temporal resolution, real-time monitoring, relatively low cost, and easy accessibility. These features make EEG an attractive technology for screening or monitoring brain-related disorders, particularly in a large number of low-income hospitals. Furthermore, EEG has found widespread use in clinical practice for providing diagnostic support for patients with neurological diseases such as epilepsy [2,3], Parkinson's disease [4], and depression [5,6]. However, EEG data often presents challenges such as non-stationary chaotic dynamics and low signal-to-noise ratio [7], which can make manual interpretation of the recorded brain signals difficult. Therefore, there is an urgent need for a reliable methodology to accurately and automatically detect EEG pathology.

Currently, several machine learning-based methodologies have been documented in the literature, including convolutional neural network [8,9], categorical boosting

(CatBoost) [10], support vector machine (SVM) [11], Riemannian geometry (RG) [7], and others. These methodologies can be broadly categorized into feature-based and deep learning approaches. The two methodologies differ in terms of classification performance and evaluation of clinical implications [12,13]. The deep learning methodologies are data-driven, which adopt an end-to-end learning framework capable of automatically extracting and classifying features. However, they require a large number of labeled samples (e.g., known patient characteristics) for model training to prevent overfitting and ensure generalization ability [14]. Additionally, deep learning models are complex, creating a trade-off between accuracy and interpretability [15]. In contrast, the feature-based methodologies can learn effective parameters with relatively small training samples [16], using two stages: EEG feature extraction using time or time-frequency analysis techniques and classification using conventional classifiers such as CatBoost and SVM. The most common methodologies for feature extraction include wavelet packet decomposition (WPD) [17], continuous wavelet transform (CWT) [7], and wavelet decomposition [18]. For instance, Zhou et al. [19] incorporated the Hilbert transform (HT) and discrete wavelet transform (DWT) for extracting amplitude and frequency modulation features in EEG data. Overall, feature-based decoding frameworks can perform as well as the most advanced deep neural networks in EEG pathology classification [7,19]. Furthermore, researchers favor these techniques due to their robustness, low cost, and interpretability [20–22]. However, the majority of existing feature-based EEG pathology methodologies primarily focus on feature extraction procedures, neglecting redundancy and irrelevant features, which may adversely impact classification performance.

In this paper, we propose a novel machine learning framework that addresses the challenges of EEG pathology detection. Our framework is based on the Kruskal–Wallis (KW) test and WPD to discriminate between normal and pathological EEG signals. The nonparametric KW test is a widely used method for feature selection, which is applied to analyze each feature for detecting statistically significant differences between normal and abnormal EEG signals. We first decompose brain signals into different wavelet coefficients using WPD and measure the statistical parameters of each selected coefficient as a feature vector. We then present a new mechanism based on feature aggregation and KW to rank the aggregated features and select the desired subset of available features, aiming to accelerate the subsequent classification procedure further. Finally, the selected features are fed into different ensemble learning classification algorithms (i.e., random forest (RF), CatBoost, and light gradient boosting machine (LightGBM)) to distinguish between pathological and non-pathological EEGs. Our experiments demonstrate that our proposed methodology outperforms the listed five representative baselines, with an F1-score of 87.60% and a G-mean of 88.60%. This indicates the feasibility and effectiveness of our proposal. Furthermore, the ablation experiment also shows that the designed feature selection mechanism can decrease the dimensionality of feature space while not impairing the characteristic quality, substantially simplifying the computational complexity of our approach. To the best of our knowledge, we are likely to be among the first to apply the KW statistical test in the EEG pathology detection domain. The major contributions of this study are listed as follows:

1. We design a novel feature selection mechanism incorporating feature aggregation and KW statistical tests to rank aggregated EEG features and eliminate meaningless features by set thresholds. In particular, the newly presented feature aggregation strategy can significantly reduce the dimension of EEG features, thereby simplifying the computational complexity of our framework.
2. We perform extensive experiments on the popular benchmark dataset. The proposed method with the CatBoost classifier outperforms existing approaches, providing a better supporting technique for EEG pathology diagnosis. We also verified that the designed feature selection proposal helps EEG detection performance improvement through rigorous ablation studies.

The remaining sections of this paper are arranged as follows. Section 2 provides a literature review of related works. Section 3 describes the experimental data and our

proposed technique in detail. Section 4 presents the experimental results and analysis. Finally, we conclude with suggestions on future work in Section 5.

2. Related Works

In recent years, machine learning has shown promising results in the medical field. However, it can be challenging for neurologists to observe important information in EEGs without utilizing machine learning technologies. Various approaches have been proposed in the literature, broadly categorized into feature-based and deep learning methods. In this section, we will briefly discuss the related works that have been employed to detect EEG pathology.

2.1. Deep Learning Methods

Deep learning approaches have shown good representation learning abilities in classification and detection tasks, and several deep learning-based approaches have been proposed for distinguishing pathological vs. normal EEG in the literature with good classification results [5,23]. These approaches adopt different neural network architectures and neural architecture search strategies to improve the performance of EEG detection. For example, Schirrmeister et al. [8] used shallow and deep convolutional neural networks with an automatic hyperparameter optimization algorithm to decode EEG pathology and obtained an accuracy of about 85%. Since then, deep learning-based EEG pathology detection approaches have proliferated. Gemein et al. [7] proposed an adaptation braindecode temporal convolutional network (BD-TCN) for optimized EEG decoding, which achieved 86.20% accuracy. Roy et al. [14] designed the ChronoNet architecture as a recurrent neural network to process EEG data efficiently and obtained 86.57% accuracy. Amin et al. [11] presented a cognitive smart healthcare framework based on the AlexNet model for analyzing EEG pathology, which resulted in 87.32% accuracy. Although these network models have shown the capability to detect abnormal EEG signals, they are complex and require a large volume of labeled samples for training to prevent overfitting and provide generalization ability [14,15]. For instance, some studies [11,24] employed massive additional closed source EEG data to pretrain the network model. Furthermore, overfitting was observed in some designed network models for EEG detection [14,25].

2.2. Feature-Based Methods

Over the past two decades, traditional machine learning technologies have been extensively utilized for EEG decoding tasks due to their robustness, low cost, and interpretability [20–22]. These machine learning approaches can be divided into two parts: feature extraction and EEG classification. They first extracted various EEG features from different domains, and then classified the extracted features into different categories. For example, Subasi et al. [2] used the DWT and WPD to extract EEG features for classifying epileptic focal region. Similarly, Acharya et al. [3] utilized the WPD to extract features and selected a subset for epileptic activity classification, while Kutlu et al. [26] used WPD for arrhythmia classification. Tawhid et al. [27] presented a completed non-parametric local transform approach for classifying multiple neurological abnormalities, and Nicolaou et al. [28] investigated permutation entropy features in the time-domain perspective for the purpose of detecting seizures.

Recently, feature-based techniques have gained attention for EEG pathology detection. For example, Gemein et al. [7] extracted different domain EEG characteristics using various analysis techniques, such as discrete Fourier transform (DFT), DWT, CWT, and HT, and achieved 85.90% accuracy with the RG classifier. Sharma et al. [18] designed an automatic framework based on wavelet decomposition and various classifiers to distinguish normal and abnormal EEGs. More recently, Albaqami et al. [17] extracted six statistical characteristics from each selected WPD coefficient, including average power, standard deviation, and skewness, and input them into the CatBoost classifier, achieving an accuracy of 87.68%. However, most existing feature-based EEG pathology detection approaches ignore

redundant and irrelevant features, which can negatively affect classification performance. Therefore, in this paper, we present a novel feature extraction mechanism to effectively address this challenge.

3. Materials and Methods

This section introduces a new framework for resolving the classification problem of normal versus pathological EEGs. The machine learning pipeline employed in this study is illustrated in Figure 1. The process involves three phases: statistical features extraction, significant features selection, and classification. Each phase is discussed in detail in the following subsections.

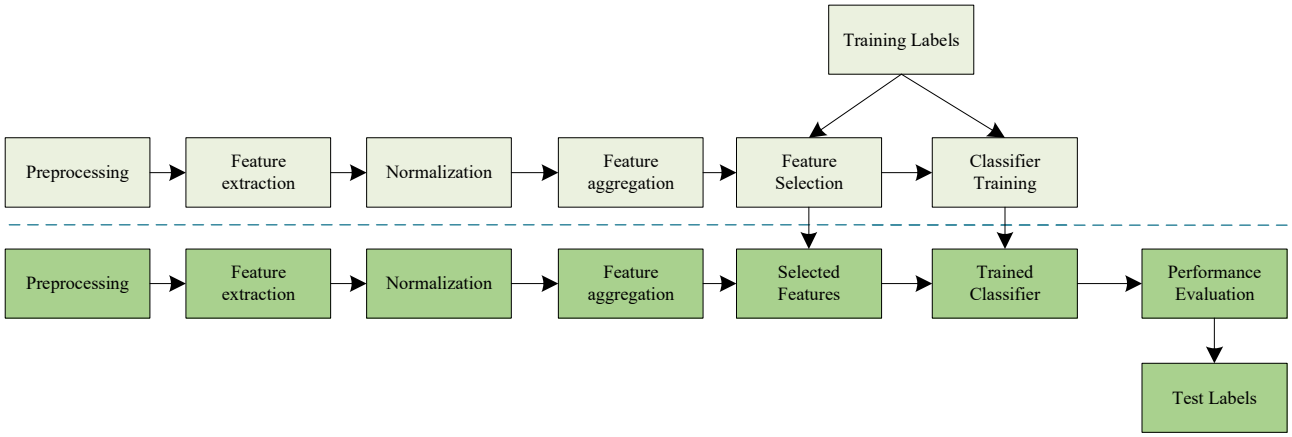

Figure 1. Flowchart of the proposed EEG pathology detection framework.

3.1. Clinical Scalp EEG Dataset

This work utilizes the publicly available EEG Abnormal Corpus V2.0.0 [29], which has been extensively studied in the literature, as in [7,8,11,30]. This database was provided by the Temple University Hospital in America and contains over 2717 multichannel EEG records collected from patients ranging in age from 7 days to 96 years. The recorded brain-related pathologies include, but are not limited to, epilepsy, stroke, and depression. The EEG recordings were obtained using common electrodes arranged according to the international 10/20 position system, as depicted in Figure 2. The database contains a minimum of 21 and a maximum of 31 EEG channels, with each recording lasting at least 15 min, and having a sampling rate ranging from 250 to 500 Hz. To ensure data consistency, the EEG signals were down-sampled to 250 Hz, and the same 21 EEG channels present in all recordings were selected (Figure 2). Additionally, neurologists have labeled each record as either pathological or normal EEG. Based on this labeling, the benchmark corpus has been separated into training and testing sets. The training set consists of 1346 pathological samples and 1371 normal samples, while the testing set consists of 126 pathological samples and 150 normal samples. The training set was obtained from 893 pathological patient records and 1237 normal patient records, while the testing set was obtained from 105 pathological patient records and 148 normal patient records. Table 1 provides a concise summary of the TUH Abnormal EEG dataset, and further information about the dataset can be found in [29].

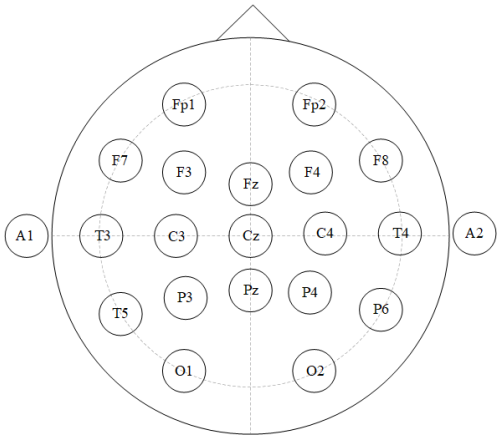

Figure 2. Layout of EEG electrodes.

Table 1. Description of samples and patients employed in TUH Abnormal EEG Corpus (V2.0.0).

Description	Samples		Patient	
	Normal	Pathological	Normal	Pathological
Training	1371	1346	1237	893
Testing	150	126	148	105
Total	1521	1472	1385	998

3.2. Feature Extraction

In general, EEG-related characteristics can effectively be extracted from raw brain data using time, frequency, and time-frequency analysis methods. However, each approach has its own strengths and limitations. The time-domain analysis techniques capture EEG features from a time-domain perspective, but may overlook valuable frequency information. The frequency-domain analysis, on the other hand, mainly focuses on the spectral structure in brain activity signals and may lose significant information. By contrast, time-frequency analysis techniques, such as WPD, can simultaneously consider both time and frequency domain information, which has gained widespread attention in recent decades and has been widely used for pathology detection [17,31,32]. Therefore, in this paper, we utilize the WPD to extract discriminative characteristics from the raw brain signals.

To address the non-stationary characteristics of long duration EEG signals, they are typically clipped into shorter segments to extract discriminative features more effectively [33,34]. Each segment is considered pseudo-stationary because of the similarity between the statistical characteristics in the time and frequency domains [35]. We introduce a non-overlapping sliding window with a size of 8 s to clip each of the EEG signals (channel-wise) into small intervals. Subsequently, the WPD technology is applied to decompose each segmented signal into frequency sub-bands. Specifically, the information obtained from the decomposition in the first level provides both approximate and detailed coefficients. WPD continues to decompose both approximate and detailed coefficients at the next level and repeats this procedure until all predetermined decomposition levels are reached. We assume W_j^n is the nth ($n = 0, 1, 2, \cdots, 2^j - 1$) wavelet packet at the jth scale, which could be represented as a wavelet function:

$$W_{j,k}^n(t) = 2^{-j/2} W^n(2^{-j}t - k), \tag{1}$$

is the orthonormal basis corresponding to W_j^n, where j, k, n, t are index scale (integer), shift factor (integer), frequency factor, and time, respectively. $W_{j,k}^n(t)$ satisfies the Equations (2) and (3) [36], when n is even,

$$W_{j,0}^n(t) = \sum_k h_0(k) W_{j-1,k}^i, \tag{2}$$

when n is odd,

$$W_{j,0}^{n}(t) = \sum_{k} h_1(k) W_{j-1,k}^{i}, \tag{3}$$

where $h_0(k), h_1(k)$ is a couple of quadruple mirror filters that are scale-independent and satisfies with (4),

$$h_1(k) = (-1)^{1-k} h_0(1-k). \tag{4}$$

Note that $W_{0,0}^{0}$ and $W_{0,0}^{1}$ represent the scaling function and the mother wavelet function, respectively. Assuming that $f(t)$ is a function of space $L^2(R)$, when the scale is small enough, the sampling sequence $f(k\Delta t)$ of $f(t)$ can be directly used to approximate the coefficients $d_0^0(k)$ as space W_0^0. The coefficient $d_j^n(k)$ of WPD at jth level and kth sample is represented as (5) and (6),

$$d_j^{2n}(k) = \sum_{k} h_0(m-2k) d_{j-1}^{n}(m), \tag{5}$$

$$d_j^{2n+1}(k) = \sum_{m}^{k} h_1(m-2k) d_{j-1}^{n}(m), \tag{6}$$

where $m \in \mathbb{Z}$. From level $(j-1)$th, we can obtain the decomposition coefficients of level jth, and the WPD coefficients of the digital signal $f(k)$ at each level can be found by sequential analogy. Figure 3 shows an example of a 3-level ($j \in \{1, 2, 3\}$) WPD decomposition of the EEG signal, where A and D stand for the approximate and detail coefficients, respectively. The effectiveness of the WPD technique is influenced by two important factors: the number of decomposition levels and the wavelet basis used. Typical wavelet basis functions include Symlets, Coiflets, and Daubechies [37]. Among these, Symlets, also referred to as orthogonal wavelets due to their good symmetry and orthogonality, is particularly well-suited for analyzing EEG data [38]. In addition, the more decomposition levels are utilized, the more coefficients are generated, resulting in a greater time overhead. In this study, in response to the results of the preliminary experiment, we choose the 8-level decomposition level and sym4 as the wavelet mother function.

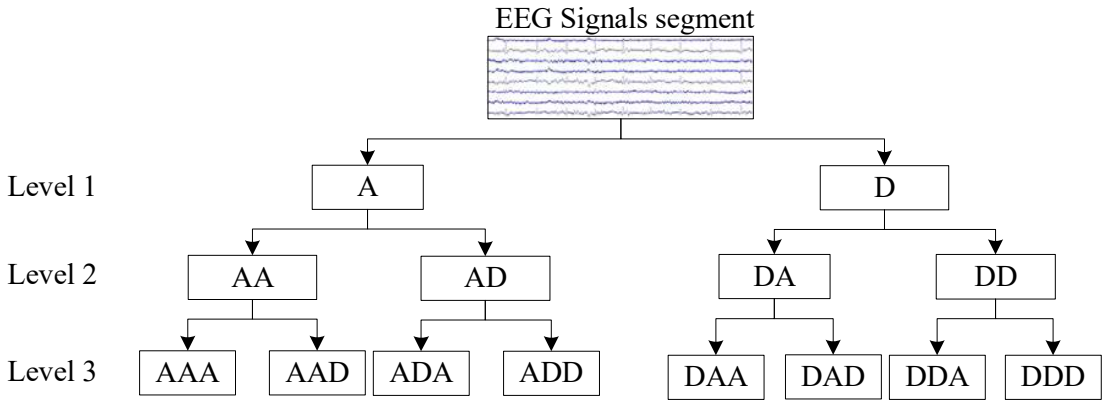

Figure 3. The structure of 3-level WPD.

After decomposition, a simple approach is to concatenate all the coefficients into a single vector and use it as the input to the classification algorithm. However, this method suffers from poor classification performance due to the high-dimensional nature and low signal-to-noise ratio of EEGs. To address this issue, a commonly adopted strategy is to derive a set of statistical characteristics from each selected frequency component, as depicted in Figure 4. In this study, we follow the approach proposed in [39,40], which involves extracting only three statistical characteristics from each selected WPD frequency component, namely, mean absolute value (m_{sk}), mean value (u_{sk}), and standard deviation (σ_{sk}). They are given by the following mathematical equations:

$$m_{sk} = \frac{1}{L_k} \sum_{t=1}^{L_k} |x_{kt}| \tag{7}$$

$$u_{sk} = \frac{1}{L_k} \sum_{t=1}^{L_k} x_{kt} \tag{8}$$

$$\sigma_{sk} = \left(\frac{1}{L_k - 1} \sum_{t=1}^{L_k} (x_{kt} - u_{sk})^2 \right) \tag{9}$$

where s represents the s-th segment of a raw signal, $k (k = 1, 2, \cdots, K)$ is the k-th coefficient in the selected coefficient list, K is the size of the selected coefficient list, L_k is the length of the k-th coefficient signal, and x_{kt} denotes the t-th data point in the k-th coefficient. Among all measures, the former two refer to the frequency information of the signal, while the σ_{sk} represents the number of changes in the frequency of the signal. For each EEG sample, after calculating the statistical characteristics, the obtained values are stored in a matrix $F_{statistics} \in R^{C \times S \times 3K}$ for further analysis:

$$F_{statistics} = \left\{ \begin{bmatrix} m_{11} & u_{11} & \sigma_{11} & \cdots & m_{1k} & \cdots & \sigma_{1K} \\ m_{21} & u_{21} & \sigma_{21} & \cdots & m_{2k} & \cdots & \sigma_{2K} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{s1} & u_{s1} & \sigma_{s1} & \cdots & m_{sk} & \cdots & \sigma_{sK} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{S1} & u_{S1} & \sigma_{S1} & \cdots & m_{Sk} & \cdots & \sigma_{SK} \end{bmatrix} \times C \; channels \right\} \tag{10}$$

where S represents the total amount of the consecutive segments for an EEG sample and C denotes the channel number of the used EEGs. In addition, for the purpose of eliminating the effects of unit and scale differences among characteristics, the statistical features are normalized with the Z-score normalization technique in this study.

Subsequently, we present the computational complexity of our WPD-based characteristics to analyze the computational cost and time requirements for computing characteristics. In this regard, we defined the big-O notation $\mathcal{O}(.)$, which generally expresses the asymptotic behavior when the input size u (here referring to the EEG signals) tends to infinity [41]. We define the time complexity T_c and g as two positive functions with the following Equation (11):

$$T_c(u) = \mathcal{O}(g(u)) \tag{11}$$

where $u \in \mathbb{R}$ if and only if positive constant t and u', such that we have $T_c(u) \leq tg(u)$ with respect to all $u \geq u'$ [17]. Let L_s be the length of the s-th ($s = 1, 2, \cdots, S$) segment, then the s-th segment time complexity of computing WPD-based characteristics can be obtained by $\mathcal{O}(L_s \log_2(L_s) * d)$, where d is the decomposition levels. It should be noted that the time complexity of feature extraction is linearly related to the length of the EEG segments. Furthermore, the total time complexity can be obtained by multiplying the time complexity of a segment by S and by C.

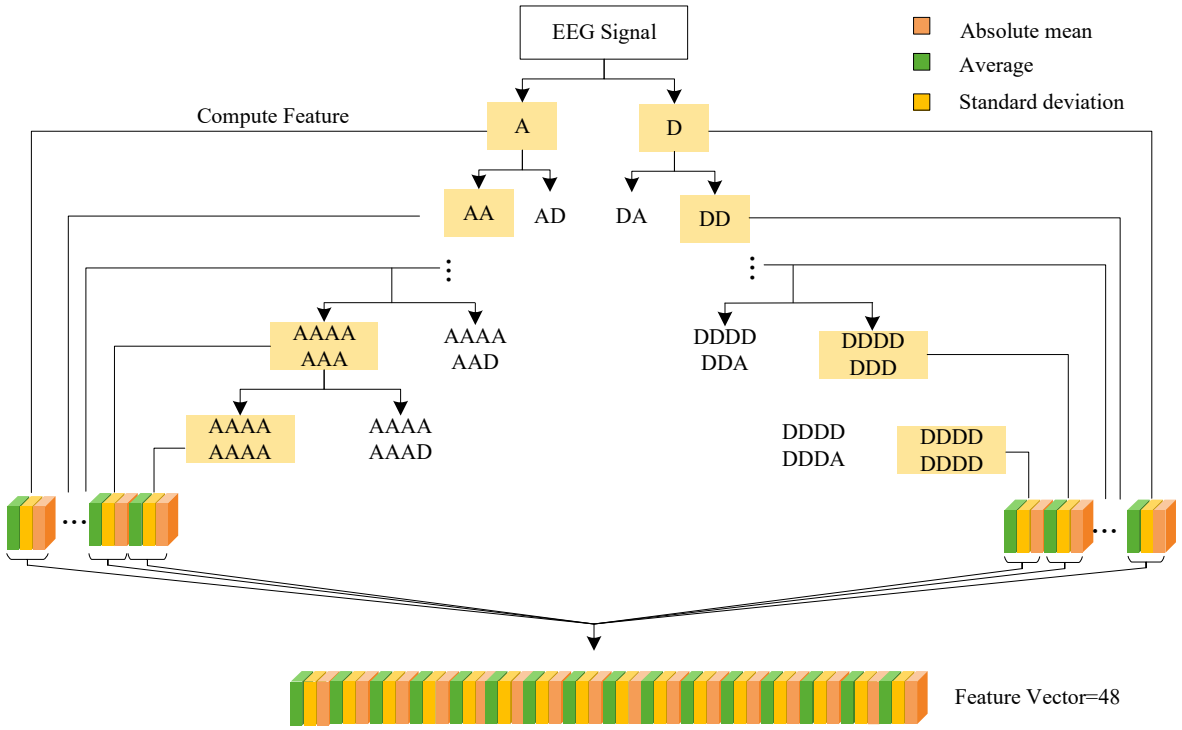

Figure 4. Statistical feature extraction process from the raw EEG signal.

3.3. Feature Selection by Statistical Significance

In the task of EEG classification, irrelevant or partially relevant extracted features can negatively affect classification performance [42]. To improve the differentiation of different categories, it is important to select more meaningful characteristics. For instance, in this study, a training set of 2717 EEG samples was used, with K, C, and S being 16, 21, and 100, respectively, for each sample. Using the proposed statistical feature extraction methodology, a one-dimensional feature vector with a length of 33,600 was derived from each sample. However, this high-dimensional feature vector increases the computational load, leading to overfitting and containing meaningless features. Therefore, appropriate feature selection is a crucial and essential step to eliminate irrelevant or redundant features. This paper proposes a novel feature selection mechanism that incorporates feature aggregation and KW statistical tests. The newly presented approach consists of two steps. In the first step, we apply a piecewise aggregation approximation technique to reduce the dimension of the feature matrix and in order to accelerate feature selection. This technique represents the original time series as a sequence, where the elements are composed of the mean values of each equal-length segment. It transforms long-time series into short ones, where the mean vector becomes a reduced representation of the data. Figure 5 depicts the entire procedure of EEG statistical feature aggregation. To minimize the negative effect on the quality of extracted characteristics, the statistical feature matrix of each EEG sample is split into the first, middle, and last three parts along with the time direction. The mean of all segments within each part is then calculated and concatenated into a single feature vector. That is, $F_{statistics}$ can be transformed into a new feature vector $\overline{F}_{statistics} = (\overline{F}_{first}, \overline{F}_{middle}, \overline{F}_{last}) \times C$, where its length is $C \times S' \times 3K$, S' denotes the number of aggregation output values for each statistical feature vector and its value is equal to 3 in this paper, that is, $S' \ll S$. $\overline{F}_{first}$ can be calculated by:

$$\overline{F}_{first} = (\overline{m}_{first,\,1}, \overline{u}_{first,\,1}, \overline{\sigma}_{first,\,1}, \cdots, \overline{\sigma}_{first,\,k}, \cdots, \overline{m}_{first,\,K}, \overline{u}_{first,\,K}, \overline{\sigma}_{first,\,K}), \quad (12)$$

where

$$\overline{m}_{first,k} = \frac{1}{y}\sum_{s=1}^{S/2} m_{sk},\ \overline{u}_{first,k} = \frac{1}{y}\sum_{s=1}^{S/2} u_{sk},\ \overline{\sigma}_{first,k} = \frac{1}{y}\sum_{s=1}^{S/2} \sigma_{sk}, y = S/2,\ k = 1, 2, \cdots, K. \quad (13)$$

Similarly, for $\overline{F}_{middle}$, $\overline{F}_{last}$, they can be obtained by:

$$\overline{m}_{middle,k} = \frac{1}{y}\sum_{s=S/4}^{3S/4} m_{sk},\ \overline{u}_{middle,k} = \frac{1}{y}\sum_{s=S/4}^{3S/4} u_{sk},\ \overline{\sigma}_{middle,k} = \frac{1}{y}\sum_{s=S/4}^{3S/4} \sigma_{sk},\ k = 1, 2, \cdots, K, \quad (14)$$

$$\overline{m}_{last,k} = \frac{1}{y}\sum_{s=S/2}^{S} m_{sk},\ \overline{u}_{last,k} = \frac{1}{y}\sum_{s=S/2}^{S} u_{sk},\ \overline{\sigma}_{last,k} = \frac{1}{y}\sum_{s=S/2}^{S} \sigma_{sk},\ k = 1, 2, \cdots, K. \quad (15)$$

In this aggregation process, a mean value is used to summarize the information embedded in $S/2$ feature values, resulting in a significant reduction in the size of the feature matrix of each part.

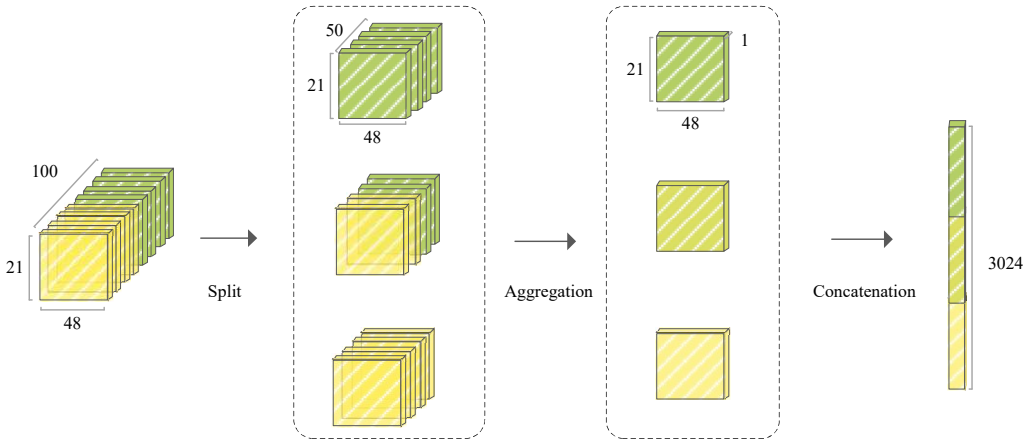

Figure 5. The complete process of feature aggregation.

In the subsequent step, an appropriate feature ranking technique is employed to further eliminate redundant or noisy characteristics that may still be present in the aggregated features. This scheme ranks all of the extracted characteristics based on their discriminative power, thereby selecting the most relevant characteristics from the available characteristics. This approach reduces the computational burden of the classifiers by using a smaller number of characteristics without compromising classification accuracy [4,43]. To analyze the significance of the aggregated characteristics, we utilize the KW test, which is a non-parametric statistical test. The highly important characteristics indicate good discrimination and can be evaluated by employing the p-value, where a smaller p-value corresponds to a more important feature. In addition, to determine the statistical significance of the aggregated features, the KW test is conducted to test the null hypothesis that there is no difference between the features of different categories. On the contrary, the alternative hypothesis is that the distribution of features between different categories is not completely the same. If the p-value is less than the selected significance level, we reject the null hypothesis. By comparing the p-value with the significance level, significant features can be distinguished, and those with a p-value less than the predetermined threshold are identified as useful features and retained, while the rest are discarded. In this work, the significance level of p is set to 0.001, which is a commonly used p-value, for example, in medical applications to perform statistical tests of significance [44,45]. More specifically, features with p-values < 0.001 are selected, indicating that there is strong evidence against the null hypothesis with less than 0.1% probability. Therefore, the calculated p-values are used to rank the features and identify an optimal set of features to reduce the computational complexity of the proposed method.

3.4. Ensemble Learning Classifiers

After selecting the strongly correlated features, we used the constructed feature matrix as the input for the classifier to perform EEG classification. Common classification algorithms include SVM, CatBoost [17], RG [7], LightGBM [46], and RF [47], among others. However, a single learner may result in poor generalization performance under the miss selection in the process of learning tasks. Therefore, ensemble learning, especially gradient boosting-based methods, has attracted significant attention due to its excellent performance and flexibility [48]. In this work, we adopted three popular ensemble learning algorithms, namely RF, LightGBM, and CatBoost, to classify the selected important characteristics. The latter two are gradient boosting-based ensemble learning classification approaches. RF is a multi-tree integration algorithm that is extensively used for classification or regression. It designates the category with the highest number of votes as the final output using the bootstrap resampling technique, while scoring the importance of the variables.

LightGBM is a light gradient boosting algorithm, also derived from gradient boosting decision trees (GBDTs) algorithms. It addresses the problem of an excessive number of samples and features three techniques, i.e., the histogram algorithm, the one-sided gradient sampling algorithm, and the mutually exclusive feature linking algorithm. These techniques make the GBDT algorithm more lightweight and further improves the training speed.

CatBoost is another novel ensemble algorithm derived from the GBDTs, proposed by Yandex [10] and demonstrated to have good ability in handling heterogeneous data. Moreover, this algorithm employs oblivious trees as basic predictors, which are balanced, to overcome over-fitting problems. Generally it exhibits superior classification performance compared to other ensemble approaches.

4. Results and Discussion

In this section, we first introduce the experimental initialization, then show the implementation details of our proposed framework, and finally conduct a series of experiments to verify the effectiveness of our proposed method. All experiments were conducted on a workstation equipped with an Intel i7-8700 processor (up to 3.20 GHz) and an NVIDIA GeForce RTX 3090 GPU.

4.1. Experimental Setup

To ensure a rigorous comparison, we used the same experimental setup as [17] to evaluate the feasibility of our proposed methodology. Specifically, we applied the same pre-processing operations, and sliced each EEG signal into 100 segments by using a non-overlap sliding window with a size of 8 s. We then decomposed each segment into a series of coefficients using the WPD with the sym4 wavelet basis and eight decomposition levels. We selected eight approximations and details from the WPD coefficients, respectively, which were exactly the same as those selected in [17]. We derived a set of statistical characteristics from each selected coefficient and used a feature selection mechanism based on the KW test to eliminate meaningless aggregated features. After carrying out the KW test, the significant characteristics with a p-value less than 0.001 were selected and fed into RF, LightGBM, and CatBoost classifiers for the detection of pathological and normal EEGs. These classifiers are briefly described in Section 3.4. In all experiments, RF, LightGBM, and CatBoost models were iterated 48, 60, and 800 times, respectively, and the maximum tree depths were set to 8, 10, and 4. The learning rates for the latter two were 0.0284 and 0.003. For the remaining hyperparameters used in each classifier, they were fixed to the default values. The details of the proposed method are shown in Table 2.

Table 2. Parameters for the proposed methodology in this work.

		Parameters	Value
Feature extraction	WPD	wavelet basis	sym4
		decomposition level	8
		selected coefficients	[D, . . . , DDDDDDDD, A, . . . , AAAAAAAA]
	KW	*p*-value	0.001
Classification methods	RF	n_estimators	48
		max_depth	8
		criterion	gini
	LightGBM	n_estimators	60
		max_depth	10
		learning_rate	0.0284
	CatBoost	n_estimators	800
		max_depth	4
		learning_rate	0.03

Moreover, we evaluated the effectiveness of the proposed framework with three commonly used measures: accuracy (ACC), F1-score, and G-mean [6,49]. Specifically, accuracy is the simplest and most intuitive evaluation metric in the classification problem, representing the proportion of correctly predicted normal and pathological samples. The F1-score is a comprehensive evaluation metric that considers both the precision and recall of the methodology, and it is not affected by sample proportion. The G-mean is obtained from the squared root of precision and recall, measuring the model's balance performance. The mathematical definitions for accuracy, G-mean, and F1-score are as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{16}$$

$$\text{G-mean} = \sqrt{\frac{TP}{TP + FN} \times \frac{TN}{TN + FP}} \tag{17}$$

$$\text{F1-score} = \frac{2TP}{2TP + FP + FN} \tag{18}$$

where TP (true positive) and TN (true negative) represent the number of correctly classified pathological and normal EEG signals. Likewise, FP and FN represent the false positive and false negative values, respectively.

4.2. Comparison with the State-of-the-Art Baselines

To validate the effectiveness of the selected attributes for EEG pathology classification, we conducted experiments with the proposed framework utilizing different classifiers on the same dataset. The experimental results are illustrated in Figure 6. The horizontal axis represents the target class, while the vertical axis represents the predicted class, with both normal and abnormal categories included. Our findings reveal that when using the CatBoost classifier, 6.6% of normal testing samples were misclassified as abnormal, and 15.8% of abnormal testing samples were misclassified as normal. Similarly, our framework achieved an accuracy of 86.95% and a G-mean of 86.40% using LightGBM, and an accuracy of 86.59% and a G-mean of 85.75% using RF. By comparison, we can conclude that the proposed framework based on CatBoost produces superior classification performance and is therefore chosen as the optimal classifier for our methodology.

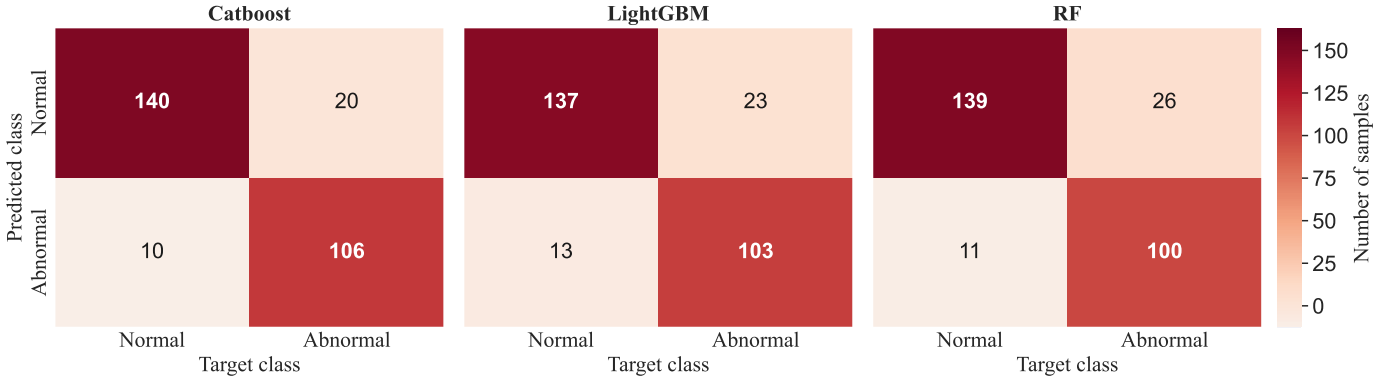

Figure 6. Confusion matrices of different classifiers on the default test set.

Next, to further validate our proposed methodology, we compared it with other state-of-the-art baselines using the same database. To ensure a fair comparison, the comparative approaches utilized the same EEG channels as our work, including BD-Deep4 [8], RG [7], AlexNet + SVM [11], WPD + CatBoost [17], and BD-TCN [7]. Table 3 presents a comparison of our methodology against five baselines, with the best results highlighted in bold.

Table 3. Comparison of results with different methods using the same public EEG dataset.

	Methods	ACC (%)	F1-Score (%)	G-Mean (%)
Traditional Models	DFT, DWT, CWT, HT + RG [7]	85.90	80.41	84.92
	WPD + CatBoost [17]	87.68	86.06	87.22
Deep Learning Models	BD-Deep4 [8]	85.42	85.42	84.06
	AlexNet + SVM [11]	87.31	84.98	86.24
	BD-TCN [7]	86.20	81.94	85.44
Proposed Method	WPD + KW + CatBoost	**89.13**	**87.60**	**88.60**

The results suggest that our approach performs favorably compared to all the baselines on the same EEG abnormal dataset. Through further analysis of Table 3, we have made the following observations: compared to the traditional approach based on the RG classifier, our approach improves classification accuracy by approximately 3.23%. In particular, our method performs better than the method presented in [17] using the same classifier, thereby providing further evidence for the effectiveness of our framework. Notably, in terms of extracted statistical features, [7,17] extracted more statistical features than our method. For example, the six statistical features (e.g., mean absolute values, skewness, kurtosis, etc.) are extracted in [17], while our approach only extracts three statistical features from each selected WPD component.

Furthermore, our approach shows improved performance compared to the two advanced deep learning approaches, which achieved an accuracy rate of 3.71% higher than the BD-Deep4 and 2.93% higher than the BD-TCN. The results suggest that the feature extraction technique proposed in this paper might be able to improve the detection of abnormal EEG signals. Similarly, our method also compared favorably with the transfer learning-based method, with a 1.82% improvement in classification accuracy and a 2.62% improvement in F1-score. It is worth mentioning that both complexities of the network structure and the number of training parameters are simultaneously increased when using deep learning methods, which is a leading cause for adopting our method to address abnormal EEG detection problems.

In conclusion, these results suggest that our proposal is effective and demonstrate the potential benefits of reducing feature redundancy in enhancing classification performance. To the best of our knowledge, feature-based approaches are better suited to deal with data scarcity, while deep learning methods involve a substantial number of labeled training

data to ensure their discriminative ability. Therefore, we chose to adopt the feature-based method to address abnormal EEG detection problems in this work.

4.3. Ablation Experiment

4.3.1. Importance of Feature Selection

The KW test is a useful tool for reducing the dimension of EEG features from 3024 to 2100, thereby simplifying the computational complexity of our framework. This section aims to fully investigate the actual impact of the KW test on the proposed EEG binary classification framework by comparing the performance of our approach with and without feature selection on a real-world abnormal EEG database. Additionally, we provide the execution time of the training phase, which includes the time required for classifier training, to further capture the importance of feature selection. Table 4 shows the experimental results achieved using the RF, LightGBM, and CatBoost classifiers, respectively. The best results in this table are indicated in bold.

Table 4. The result of ablation experiment.

	Feature Selection	ACC (%)	F1-Score (%)	G-Mean (%)	Training Time (s)
RF	No	84.42	81.54	83.28	4.15
	Yes	86.59	84.38	85.75	3.19
LightGBM	No	83.69	81.17	82.92	5.83
	Yes	86.95	85.12	86.40	3.68
CatBoost	No	85.86	83.54	85.02	41.55
	Yes	**89.13**	**87.60**	**88.60**	31.95

Table 4 indicates that the adopted feature selection method can significantly enhance the classification performance of our proposal. Specifically, the detection accuracy of our methodology improves from 85.86% to 89.13% after performing the KW statistical test to select significant features. Similar comparison results can also be seen for the RF and LightGBM classifiers. Furthermore, the adopted feature selection methodology can decrease the time overhead significantly. For example, the execution time for the CatBoost classifier is reduced by 9.6(s) when feature selection is employed compared to when it is not. Although the execution time spent using feature selection is more than the other two classifiers, the better classification produced by the CatBoost classifier is worth the extra time. The results suggest that the KW test-based feature selection scheme is effective and suitable for improving EEG pathology detection performance.

4.3.2. Effect of Wavelet Function and Decomposition Level on Performance

As EEG signals are known to be non-linear and non-stationary, the use of the discrete mother wavelet function with orthogonality can help preserve the decomposed EEG signal [50]. Popular choices include Daubechies, Symlets, and Coiflets. However, there is no well-established method to select a specific mother wavelet basis function that is more suitable for EEG analysis [37]. To further study the effect of the mother wavelet and decomposition level on the EEG classification performance, we experimented with 74 wavelets from orthogonal families, namely Daubeches (db1-db38), Symlets (sym2-sym20), and Coiflets (coif 1-coif 17), with the decomposition levels ranging from 4 to 8 with an interval of 2. These mother wavelets are commonly used in biomedical signal processing [51]. The CatBoost was employed to evaluate the effect of different wavelet coefficients on abnormality detection. The top five mother wavelets that exhibit optimal performance among the three wavelet families are presented in Table 5. The table's best results are highlighted in bold.

Table 5. The top five mother wavelets that exhibit optimal performance among the three wavelet families.

Mother Wavelet	4-Level		6-Level		8-Level	
	ACC (%)	F1-Score (%)	ACC (%)	F1-Score (%)	ACC (%)	F1-Score (%)
Sym3	77.17	74.28	80.79	78.18	86.23	83.89
Sym4	77.17	74.28	81.52	79.18	**89.13**	**87.60**
Sym5	77.17	74.49	82.97	80.97	86.23	84.29
db3	77.17	74.28	80.79	78.18	86.23	83.89
db4	77.17	74.28	81.52	78.83	85.50	82.90

As shown in Table 5, the optimal performance is achieved by using the Sym4 wavelet function with 8-level decomposition, surpassing the performance of the other two levels. This configuration resulted in a classification accuracy of 89.13% and an F1-score of 87.60%, indicating that 8-level decomposition is effective for EEG pathology detection. Conversely, we also observed that using too low a level of decomposition may not represent the signal information in sufficient detail, leading to unsatisfactory results.

5. Conclusions

In this work, we propose a feature-based framework for binary classification of EEG signals using the WPD and KW tests, aiming to provide a better supporting technique for pathology diagnosis in EEGs. Compared with most existing feature-based approaches, our approach extracts fewer characteristics from each selected WPD component, and meanwhile, a feature reduction technique is presented to further eliminate redundant and meaningless characteristics, which can enhance the performance of EEG pathology classification. To the best of our knowledge, we have not found any similar approach being used for EEG abnormality detection in the literature surveyed in this area. The experiments on the real EEG data suggest that our proposed framework with the CatBoost classifier performs favorably compared to other competing techniques, indicating the potential effectiveness of the proposed approach. Furthermore, the ablation experiment provides supporting evidence for the benefits of the adopted feature selection mechanism in improving classification performance.

However, we recognize that the proposed approach must be validated in real-world clinical settings, and further studies should be conducted in collaboration with neurologists and other medical professionals to ensure that the classification approach conforms to established diagnostic criteria. Furthermore, there are several meaningful directions for future research. One direction is to investigate a feasible technique for assessing the quality of the final selected features to further remove redundant information. Another direction is to extend our approach to reduce redundant EEG channels, thereby improving the efficiency of EEG analysis.

Author Contributions: Conceptualization, L.C. and T.W.; methodology, Y.Z. and H.W.; software, T.W.; validation, H.W. and T.W.; formal analysis, Y.Z. and T.W.; investigation, L.C.; resources, L.C. and T.W.; data curation, Y.Z. and H.W.; writing—original draft preparation, Y.Z. and H.W.; writing—review and editing, Y.Z. and H.W.; visualization, T.W.; supervision, L.C. and T.W.; project administration, L.C.; funding acquisition, L.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China under Grant No. 61672157, and the National Key Research and Development Program of China (No.2020YFF0401865, No.2021YFF1200700).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank all the anonymous reviewers for their insightful comments and constructive suggestions that have obviously upgraded the quality of this manuscript. Meanwhile, we would also like to gratefully thank Xiu Cheng, the attending doctor of the Electrophysiology Department of Fujian Maternal and Child Health Hospital, China, for providing kind support and help in validating the results of our experiments.

Conflicts of Interest: The authors declare that they have no known competing financial interest or personal circumstances that could have appeared to influence the work reported in this manuscript.

References

1. Schomer, D.L.; Da Silva, F.L. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*; Lippincott Williams and Wilkins: Philadelphia, PA, USA, 2012. [CrossRef]
2. Subasi, A.; Kevric, J.; Abdullah, C.M. Epileptic seizure detection using hybrid machine learning methods. *Neural Comput. Appl.* **2019**, *31*, 317–325. [CrossRef]
3. Acharya, U.R.; Sree, S.V.; Alvin, A.P.C.; Suri, J.S. Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework. *Expert Syst. Appl.* **2012**, *39*, 9072–9078. [CrossRef]
4. Chawla, P.; Rana, S.B.; Kaur, H.; Singh, K.; Yuvaraj, R.; Murugappan, M. A decision support system for automated diagnosis of Parkinson's disease from EEG using FAWT and entropy features. *Biomed. Signal Process. Control* **2023**, *79*, 104116. [CrossRef]
5. Acharya, U.R.; Oh, S.L.; Hagiwara, Y.; Tan, J.H.; Adeli, H.; Subha, D.P. Automated EEG-based screening of depression using deep convolutional neural network. *Comput. Methods Programs Biomed.* **2018**, *161*, 103–113. [CrossRef]
6. Khosla, A.; Khandnor, P.; Chand, T. Automated diagnosis of depression from EEG signals using traditional and deep learning approaches: A comparative analysis. *Biocybern. Biomed. Eng.* **2022**, *42*, 108–142. [CrossRef]
7. Gemein, L.A.W.; Schirrmeister, R.T.; Chrabąszcz, P.; Wilson, D.; Boedecker, J.; Schulze-Bonhage, A.; Hutter, F.; Ball, T. Machine-learning-based diagnostics of EEG pathology. *NeuroImage* **2020**, *220*, 117021. [CrossRef]
8. Schirrmeister, R.T.; Gemein, L.; Eggensperger, K.; Hutter, F. Deep learning with convolutional neural networks for decoding and visualization of eeg pathology. *arXiv* **2017**, arXiv:1708.08012.
9. Lai, C.Q.; Ibrahim, H.; Suandi, S.A.; Abdullah, M.Z. Convolutional Neural Network for Closed-Set Identification from Resting State Electroencephalography. *Mathematics* **2022**, *10*, 3442. [CrossRef]
10. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. In Proceedings of the Advances in Neural Information Processing Systems 31 (NeurIPS 2018), Montreal, QC, Canada, 3–8 December 2018.
11. Amin, S.U.; Hossain, M.S.; Muhammad, G.; Alhussein, M.; Rahman, M.A. Cognitive Smart Healthcare for Pathology Detection and Monitoring. *IEEE Access* **2019**, *7*, 10745–10753. [CrossRef]
12. Korotcov, A.; Tkachenko, V.; Russo, D.P.; Ekins, S. Comparison of deep learning with multiple machine learning methods and metrics using diverse drug discovery data sets. *Mol. Pharm.* **2017**, *14*, 4462–4475. [CrossRef]
13. Tang, J.; Liu, R.; Zhang, Y.L.; Liu, M.Z.; Hu, Y.F.; Shao, M.J.; Zhu, L.J.; Xin, H.W.; Feng, G.W.; Shang, W.J.; et al. Application of machine-learning models to predict tacrolimus stable dose in renal transplant recipients. *Sci. Rep.* **2017**, *7*, 42192. [CrossRef] [PubMed]
14. Roy, S.; Kiral-Kornek, I.; Harrer, S. ChronoNet: A deep recurrent neural network for abnormal EEG identification. In *Artificial Intelligence in Medicine: 17th Conference on Artificial Intelligence in Medicine, AIME 2019, Poznan, Poland, 26–29 June 2019*; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 47–56. [CrossRef]
15. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In Proceedings of the Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
16. Zhang, J.; Xiang, Y.; Wang, Y.; Zhou, W.; Xiang, Y.; Guan, Y. Network traffic classification using correlation information. *IEEE Trans. Parallel Distrib. Syst.* **2013**, *24*, 104–117. [CrossRef]
17. Albaqami, H.; Hassan, G.M.; Subasi, A.; Datta, A. Automatic detection of abnormal EEG signals using wavelet feature extraction and gradient boosting decision tree. *Biomed. Signal Process. Control* **2021**, *70*, 102957. [CrossRef]
18. Sharma, M.; Patel, S.; Acharya, U.R. Automated detection of abnormal EEG signals using localized wavelet filter banks. *Pattern Recognit. Lett.* **2020**, *133*, 188–194. [CrossRef]
19. Zhou, J.; Meng, M.; Gao, Y.; Ma, Y.; Zhang, Q. Classification of Motor Imagery EEG using Wavelet Envelope Analysis and LSTM Networks. In Proceedings of the Chinese Control and Decision Conference (CCDC), Shenyang, China, 9–11 June 2018; pp. 9–11. [CrossRef]
20. Lotte, F.; Congedo, M.; Lécuyer, A.; Lamarche, F.; Arnaldi, B. A review of classification algorithms for EEG-based brain–computer interfaces. *J. Neural Eng.* **2007**, *4*, R1. [CrossRef]
21. Kia, S.M.; Pedregosa, F.; Blumenthal, A.; Passerini, A. Group-level spatio-temporal pattern recovery in MEG decoding using multi-task joint feature learning. *J. Neurosci. Methods* **2017**, *285*, 97–108. [CrossRef]
22. Lin, Q.; Ye, S.; Wu, C.; Gu, W.; Wang, J.; Zhang, H.L.; Xue, Y. A novel framework based on biclustering for automatic epileptic seizure detection. *Int. J. Mach. Learn. Cybern.* **2019**, *10*, 311–323. [CrossRef]

23. Xiang, J.; Xu, G.; Ma, C.; Hou, J. End-to-end learning deep CRF models for multi-object tracking deep CRF models. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *31*, 275–288. [CrossRef]
24. Alhussein, M.; Muhammad, G.; Hossain, M.S. EEG pathology detection based on deep learning. *IEEE Access* **2019**, *7*, 27781–27788. [CrossRef]
25. Roy, S.; Kiral-Kornek, I.; Harrer, S. Deep learning enabled automatic abnormal EEG identification. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 17–21 July 2018; pp. 2756–2759. [CrossRef]
26. Kutlu, Y.; Kuntalp, D. Feature extraction for ECG heartbeats using higher order statistics of WPD coefficients. *Comput. Methods Programs Biomed.* **2012**, *105*, 257–267. [CrossRef]
27. Tawhid, M.N.A.; Siuly, S.; Wang, K.; Wang, H. Textural feature based intelligent approach for neurological abnormality detection from brain signal data. *PLoS ONE* **2022**, *17*, e0277555. [CrossRef] [PubMed]
28. Nicolaou, N.; Georgiou, J. Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. *Expert Syst. Appl.* **2012**, *39*, 202–209. [CrossRef]
29. López, S.; Obeid, I.; Picone, J. Automated Interpretation of Abnormal Adult Electroencephalograms. Ph.D. Thesis, College of Engineering Temple University, Philadelphia, PA, USA, 2017.
30. Nahmias, D.O.; Civillico, E.F.; Kontson, K.L. Deep learning and feature based medication classifications from EEG in a large clinical data set. *Sci. Rep.* **2020**, *10*, 14206. [CrossRef] [PubMed]
31. Göksu, H. EEG based epileptiform pattern recognition inside and outside the seizure states. *Biomed. Signal Process. Control* **2018**, *43*, 204–215. [CrossRef]
32. Lu, N.; Li, T.; Ren, X.; Miao, H. A deep learning scheme for motor imagery classification based on restricted Boltzmann machines. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2016**, *25*, 566–576. [CrossRef]
33. Anuragi, A.; Sisodia, D.S.; Pachori, R.B. Automated FBSE-EWT based learning framework for detection of epileptic seizures using time-segmented EEG signals. *Comput. Biol. Med.* **2021**, *136*, 104708. [CrossRef]
34. Azami, H.; Escudero, J. Amplitude-aware permutation entropy: Illustration in spike detection and signal segmentation. *Comput. Methods Programs Biomed.* **2016**, *128*, 40–51. [CrossRef]
35. Singh, K.; Malhotra, J. Deep learning based smart health monitoring for automated prediction of epileptic seizures using spectral analysis of scalp EEG. *Phys. Eng. Sci. Med.* **2021**, *44*, 1161–1173. [CrossRef]
36. Wu, T.; Yan, G.Z.; Yang, B.H.; Sun, H. EEG feature extraction based on wavelet packet decomposition for brain computer interface. *Measurement* **2008**, *41*, 618–625. [CrossRef]
37. Al-Qazzaz, N.K.; Hamid Bin Mohd Ali, S.; Ahmad, S.A.; Islam, M.S.; Escudero, J. Selection of mother wavelet functions for multi-channel EEG signal analysis during a working memory task. *Sensors* **2015**, *15*, 29015–29035. [CrossRef]
38. Chakrabarti, S.; Swetapadma, A.; Ranjan, A.; Pattnaik, P.K. Time domain implementation of pediatric epileptic seizure detection system for enhancing the performance of detection and easy monitoring of pediatric patients. *Biomed. Signal Process. Control* **2020**, *59*, 101930. [CrossRef]
39. Ahammad, N.; Fathima, T.; Joseph, P. Detection of epileptic seizure event and onset using EEG. *BioMed Res. Int.* **2014**, *2014*, 450573. [CrossRef] [PubMed]
40. Kusumastuti, R.D.; Wibawa, A.D.; Purnomo, M.H. Stroke Severity Classification based on EEG Statistical Features. In Proceedings of the 2021 1st International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS), Virtually, 15–16 October 2021; pp. 138–142.
41. Anari, S.; Tataei Sarshar, N.; Mahjoori, N.; Dorosti, S.; Rezaie, A. Review of deep learning approaches for thyroid cancer diagnosis. *Math. Probl. Eng.* **2022**, *2022*, 5052435.
42. Robnik-Šikonja, M.; Kononenko, I. Theoretical and empirical analysis of ReliefF and RReliefF. *Mach. Learn.* **2003**, *53*, 23–69. [CrossRef]
43. Gupta, V.; Pachori, R.B. Epileptic seizure identification using entropy of FBSE based EEG rhythms. *Biomed. Signal Process. Control* **2019**, *53*, 101569. [CrossRef]
44. Lee, J.; Yang, J.H. Analysis of driver's EEG given take-over alarm in SAE level 3 automated driving in a simulated environment. *Int. J. Automot. Technol.* **2020**, *21*, 719–728. [CrossRef]
45. Marquetand, J.; Vannoni, S.; Carboni, M.; Li Hegner, Y.; Stier, C.; Braun, C.; Focke, N.K. Reliability of magnetoencephalography and high-density electroencephalography resting-state functional connectivity metrics. *Brain Connect.* **2019**, *9*, 539–553. [CrossRef] [PubMed]
46. Ke, G.; Meng, Q.; Finley, T.; Wang, T.F.; Chen, W.; Ma, W.D.; Ye, Q.W.; Liu, T.Y. Lightgbm: A highly efficient gradient boosting decision tree. In Proceedings of the Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
47. Li, X.; Hu, B.; Sun, S.; Cai, H. EEG-based mild depressive detection using feature selection methods and classifiers. *Comput. Methods Programs Biomed.* **2016**, *136*, 151–161. [CrossRef]
48. Samat, A.; Li, E.; Du, P.; Liu, S.; Miao, Z.; Zhang, W. CatBoost for RS Image Classification with Pseudo Label Support From Neighbor Patches-Based Clustering. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]
49. Hemachandira, V.S.; Viswanathan, R. A Framework on Performance Analysis of Mathematical Model-Based Classifiers in Detection of Epileptic Seizure from EEG Signals with Efficient Feature Selection. *J. Healthc. Eng.* **2022**, *2022*, 7654666. [CrossRef]

50. Santoso, S.; Powers, E.J.; Grady, W.M.; Hofmann, P. Power quality assessment via wavelet transform analysis. *IEEE Trans. Power Deliv.* **1996**, *11*, 924–930. [CrossRef]
51. Khanam, R.; Ahmad, S.N. Selection of wavelets for evaluating SNR, PRD and CR of ECG signal. *Int. J. Eng. Sci. Innov. Technol.* **2013**, *2*, 112–119.