

Research Article

Recognition of Imbalanced Epileptic EEG Signals by a Graph-Based Extreme Learning Machine

Jie Zhou ¹, Xiongtao Zhang ², and Zhibin Jiang ¹

¹Department of Computer Science and Engineering, Shaoxing University, Shaoxing, China

²School of Information Engineering, Huzhou University, Huzhou, China

Correspondence should be addressed to Jie Zhou; 799489588@qq.com

Received 8 April 2021; Revised 29 April 2021; Accepted 28 May 2021; Published 14 June 2021

Academic Editor: Shan Zhong

Copyright © 2021 Jie Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Epileptic EEG signal recognition is an important method for epilepsy detection. In essence, epileptic EEG signal recognition is a typical imbalanced classification task. However, traditional machine learning methods used for imbalanced epileptic EEG signal recognition face many challenges: (1) traditional machine learning methods often ignore the imbalance of epileptic EEG signals, which leads to misclassification of positive samples and may cause serious consequences and (2) the existing imbalanced classification methods ignore the interrelationship between samples, resulting in poor classification performance. To overcome these challenges, a graph-based extreme learning machine method (G-ELM) is proposed for imbalanced epileptic EEG signal recognition. The proposed method uses graph theory to construct a relationship graph of samples according to data distribution. Then, a model combining the relationship graph and ELM is constructed; it inherits the rapid learning and good generalization capabilities of ELM and improves the classification performance. Experiments on a real imbalanced epileptic EEG dataset demonstrated the effectiveness and applicability of the proposed method.

1. Introduction

Epilepsy is a common neurological disease that can cause recurrent seizures. During seizures, injury or life-threatening events may occur owing to the distraction or involuntary spasms of the patient [1, 2]. In the clinical diagnosis of various seizures, electroencephalogram (EEG) signal detection plays a crucial role [3]. This is because the epileptic brain releases characteristic waves during seizures. In recent years, an increasing number of machine learning-based methods have been applied for epileptic EEG signal recognition [4–8]. Figure 1 illustrates a machine learning method-based system for epileptic EEG signal recognition. The figure shows that an epileptic EEG signal recognition system involves the following three main steps: (1) a feature extraction method is used on original epileptic EEG signals for training and testing, (2) EEG signals after feature extraction for training are used to train the machine learning-based model to build an epileptic EEG signal recognition system, and (3) EEG signals after feature extraction for testing are then inputted into the epileptic EEG signal recognition system for detection.

Previously, many machine learning methods have been proposed for epileptic EEG signal recognition, such as the naive Bayes method (NB) [9], *K*-nearest neighbor (KNN) [10], support vector machine (SVM) [11], fuzzy system [12, 13], and extreme learning machine (ELM) [14, 15], and they have shown good effectiveness. In essence, epileptic EEG signal recognition is a typical imbalanced classification task [16, 17]. Compared with negative samples (people without epilepsy), positive samples (patients with epilepsy) have extremely low representation and cannot be well classified by traditional classifiers. Although the misclassification of positive samples has little effect on the model accuracy, it may cause serious medical malpractice. Therefore, traditional machine learning methods face several critical challenges for recognition of imbalanced epileptic EEG signals: (1) traditional machine learning methods often ignore the imbalance of epileptic EEG signals and misclassify positive samples, which may cause serious medical malpractice, and (2) existing imbalanced classification methods ignore the interrelationship between samples, resulting in poor classification performance. Therefore, building a classifier that considers

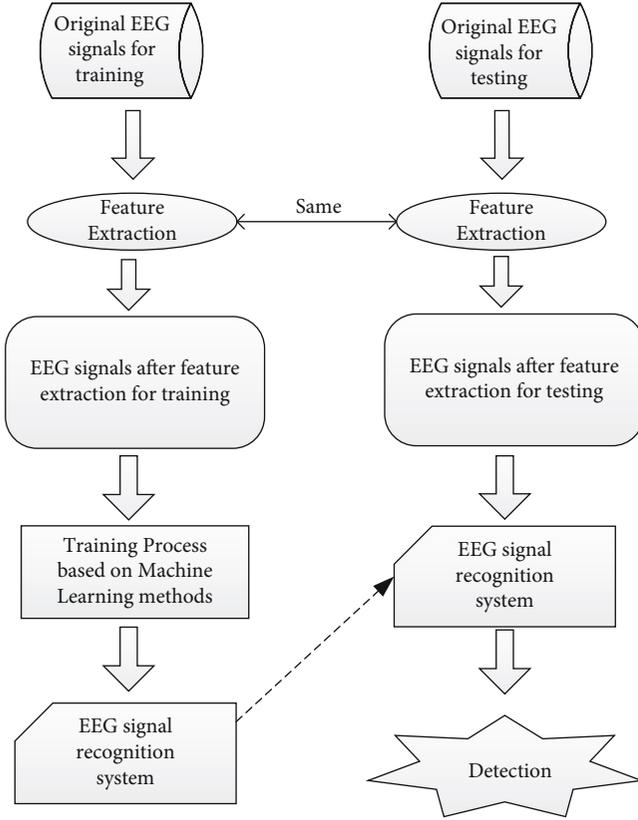


FIGURE 1: Illustration of the machine learning method-based system for epileptic EEG signal recognition.

the imbalance of the epileptic EEG signals and additional knowledge of samples becomes imperative for classification of imbalanced datasets with epileptic EEG signals.

To overcome these challenges, a novel imbalanced epileptic EEG signal recognition method based on a graph and ELM is proposed in this study. ELM has become a classical machine learning method with its solid theoretical foundations, fast training speed, and good predictive performance [18, 19]. Although ELM can universally approximate to any continuous functions, it is not effective for classifying imbalanced datasets. Therefore, it is necessary to adopt strategies to make ELM correctly classify positive samples to obtain a reasonable classification result of an imbalanced dataset. Previously, numerous imbalanced ELM-based methods have been proposed. For example, Zong et al. [20] proposed the weighted extreme learning machine (WELM), which pioneered the application of ELM in imbalanced classification. Similarly, Zhang and Ji [21] proposed a fuzzy ELM (FELM), which regulated the distributions of penalty factors by inserting a fuzzy matrix. Yu et al. [22] proposed a special cost-sensitive ELM (ODOC-ELM) for imbalanced classification problems. Li et al. [23] proposed an ensemble WELM algorithm based on the AdaBoost framework to learn the weights of different samples adaptively. Yang et al. [24] proposed a novel ELM-based imbalanced classification method by estimating the probability density distributions for two imbalanced classes. Shukla and Yadav [25] combined CC-ELM with WELM to propose a regularized weighted CC-ELM.

Xiao et al. [26] proposed an imbalanced ELM-based algorithm for two classes of classification tasks by solving each class classification error. Du et al. [27] proposed an online sequential extreme learning machine with under- and over-sampling (OSELM-UO) for online imbalanced big data classification. In addition, some ELM-based imbalanced methods, such as ensemble weighted ELM [28], class-specific cost regulation ELM [29], label-weighted extreme learning machine [30], and class-specific ELM [31], have also been proposed. However, to the best of our knowledge, there is no study that uses imbalanced ELM methods for epileptic EEG signal recognition; therefore, it is necessary to propose such a method for epileptic EEG signal recognition.

In this study, inspired by WELM, we propose a novel graph-based ELM (G-ELM) for imbalanced epileptic EEG signal recognition. First, we use the graph theory to construct a relationship graph of samples according to their data distribution. Then, we combine the relationship graph with ELM to propose G-ELM. The experimental results on a real imbalanced epileptic EEG dataset show that the proposed method can address imbalanced classification of epileptic EEG signals effectively. The main contributions of this study are as follows.

- (1) The proposed G-ELM sets the compensation for loss of positive samples to be greater than that of negative samples based on graph theory and then combines with the ELM to classify imbalanced data effectively. It is a novel imbalanced ELM-based method, which attains a good classification performance and inherits the rapid learning and good generalization capabilities of ELM
- (2) The proposed imbalanced classification method attempts to consider both the imbalance and interrelationship of epileptic EEG samples to obtain better performance for imbalanced epileptic EEG signal recognition. It can be utilized for imbalanced epileptic EEG signal recognition. It not only realizes effective classification of imbalanced epileptic EEG signals from a new perspective but also expands application of ELM-based algorithms
- (3) We use six imbalanced classification evaluation indices, i.e., accuracy, precision, recall, F -measure, G -means, and AUC, to compare the performance of the proposed G-ELM and the existing imbalanced ELM-based methods. Extensive experiments on a real imbalanced epileptic EEG dataset indicate that the proposed method can address imbalanced epileptic EEG signal recognition effectively and outperform the existing imbalanced ELM-based methods

The rest of this paper is organized as follows. Section 2 introduces the background underlying the proposed epileptic EEG recognition method. In Section 3, the details of the proposed G-ELM are presented. The performance of the proposed method is evaluated with several comparative methods in Section 4. The conclusions of this paper are provided in Section 5.

TABLE 1: Descriptions of the five groups in *Bonn*.

	Groups	Sizes of groups	Descriptions of EEG signals
<i>Healthy people</i>	A	100	EEG signals obtained with eyes open
	B	100	EEG signals obtained with eyes closed
	C	100	EEG signals measured in the hippocampus of the brain during seizure-free intervals
<i>Epileptic people</i>	D	100	EEG signals measured in the epileptogenic zone during seizure-free intervals
	E	100	EEG signals measured during seizure activity

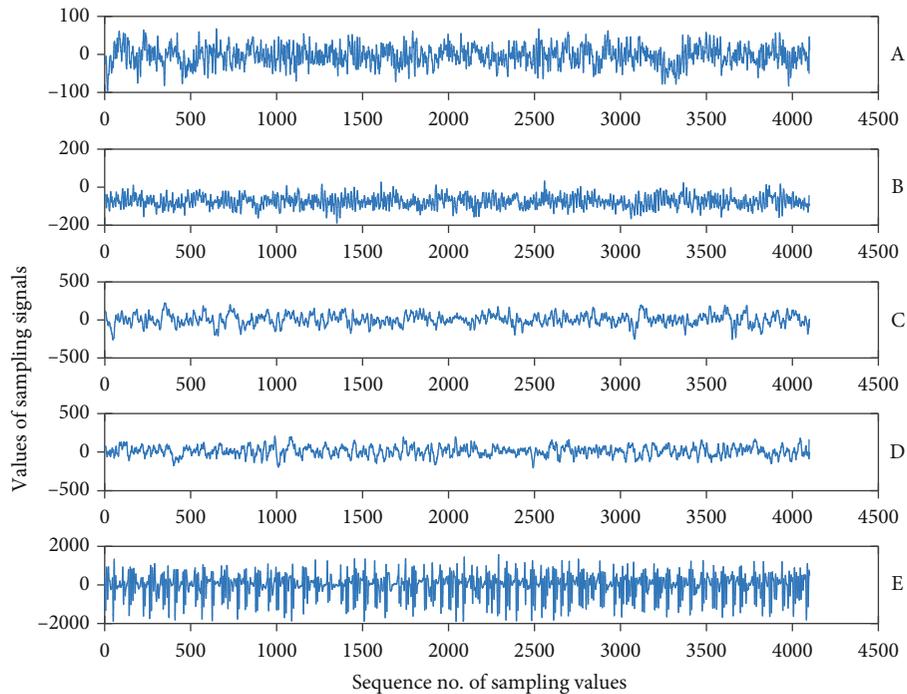


FIGURE 2: Representative original epileptic EEG signals of five different groups.

2. Background

In this section, we briefly describe the background related to the proposed epileptic EEG signal recognition method. It includes the epileptic EEG dataset, the feature extraction methods, and the classical ELM, which are used for epileptic EEG signal detection.

2.1. Epileptic EEG Dataset. The real epileptic EEG dataset used in this paper is *Bonn* [32], which is from the University of Bonn, Germany. It can be publicly downloaded from the following website (<http://www.epileptologie-bonn.de/cms/upload/workgroup/lehnertz/eegdata.html>). There are five groups (denoted by A–E, respectively) in *Bonn*. In each group, there are 100 samples of 23.6s segments. Detailed descriptions of the five groups are given in Table 1. Groups A and B are segments acquired from five healthy volunteers with eyes open (Group A) and eyes closed (Group B). Groups C–E are segments acquired from volunteers with epilepsy. In Group C, EEG signals are measured in the hippocampus of the brain during seizure-free intervals and those in Group D are measured in the epileptogenic zone during seizure-free intervals. In Group E, EEG signals are measured during

seizure activity. Five representative original epileptic EEG signals of five different groups are shown in Figure 2.

2.2. Feature Extraction. Many studies [33–35] have shown that the original EEG signals cannot be directly used for training machine learning-based models and that feature extraction is a necessary step. This is because the original EEG signals are usually high dimensional, stochastic, non-stationary, and nonlinear and the background noise in the original signals is very complex. The commonly used feature extraction methods can be divided into three main categories: time domain analysis, frequency domain analysis, and time-frequency analysis. Time domain analysis-based methods extract the features by analyzing the characteristics of original EEG signals, such as mean, variance, amplitude, and kurtosis [36]. Frequency domain analysis-based methods usually analyze the EEG signals in the frequency domain to extract the features, such as fast Fourier transforms [37] and short-time Fourier transforms [38]. As for time-frequency analysis methods, the information of time and frequency domain is considered simultaneously to extract the features from original epileptic EEG signals. Typical time-frequency analysis-based methods are wavelet transform

methods [39, 40]. In this paper, we use the wavelet packet decomposition [40] for feature extraction from original epileptic EEG signals to simultaneously utilize the information of time and frequency domain.

2.3. ELM. ELM [19], which was first proposed by Huang et al., is a single-hidden-layer feedforward neural network [41]. It can directly optimize the output weight of the hidden layer by setting the number of hidden nodes, without paying attention to the weight and offset of the input layer, which can be generated randomly. Compared with other traditional supervised learning methods, it has good generalization ability and high learning speed. Figure 3 shows the network structure of an ELM.

ELM considers both empirical and structural risks, and its objective function is as follows:

$$\begin{aligned} \min \quad & \frac{1}{2} \|\boldsymbol{\beta}\|^2 + \frac{1}{2} C \sum_{i=1}^n \|\varepsilon_i\|^2, \\ \text{s.t.} \quad & \mathbf{h}(\mathbf{x}_i)\boldsymbol{\beta} = y_i + \varepsilon_i, \quad i = 1, 2, \dots, n, \end{aligned} \quad (1)$$

where

$$\mathbf{H}_{n \times m} = (\mathbf{h}(\mathbf{x}_1), \mathbf{h}(\mathbf{x}_2), \dots, \mathbf{h}(\mathbf{x}_n))^T = \begin{pmatrix} h_1(\mathbf{x}_1) & \dots & h_m(\mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ h_1(\mathbf{x}_n) & \dots & h_m(\mathbf{x}_n) \end{pmatrix} \quad (2)$$

represents the hidden layer feature matrix, where $h_i(\mathbf{x}_j) = g(\mathbf{A}^{(i)}\mathbf{x}_j + b_i)$, $\mathbf{A}^{(i)}$ represents the i th row of the weight matrix $\mathbf{A}_{m \times d}$, $\mathbf{b} = (b_1, b_2, \dots, b_m)^T$ represents the offset, $\mathbf{x}_i \in \mathbb{R}^d$, $i = 1, 2, \dots, n$ denotes the training samples, n is the number of training samples, d is dimension, and m is the number of hidden nodes; $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$ is the error matrix between the network outputs and the target outputs. C is a penalty parameter, which can adjust the accuracy and generalization ability of the ELM.

The optimization problem in (1) can be solved based on the Karush–Kuhn–Tucker theory. The output weight of ELM can be calculated by

$$\boldsymbol{\beta} = \begin{cases} \mathbf{H}^T \left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{Y} & n < m, \\ \left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{H}^T \mathbf{Y} & n \geq m. \end{cases} \quad (3)$$

3. Graph-Based Extreme Learning Machine

In this section, a graph-based ELM (G-ELM) is proposed. We first introduce the relationship graph of an imbalanced dataset and then develop the proposed imbalanced classification method G-ELM by combining the relationship graph with an ELM.

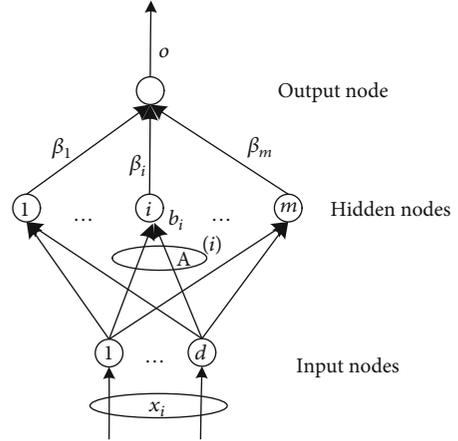


FIGURE 3: ELM network structure.

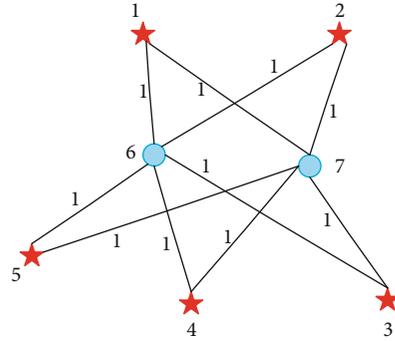


FIGURE 4: Graph structure of an imbalanced synthetic dataset.

3.1. Relationship Graph of an Imbalanced Dataset. In the context of imbalanced classification problem, the relationships between the training samples can be regarded as an undirected graph.

Undirected graph can be expressed as $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, where \mathbf{V} is the vertex set of graph \mathbf{G} and \mathbf{E} is the edge set of graph \mathbf{G} . Figure 4 shows an undirected graph of an imbalanced synthetic dataset with 7 samples, where 2 positive samples are represented by a *blue circle* and 5 negative samples are represented by a *red star*. All samples are numbered for subsequent display. Note that there are connections between samples in different classes and the weight is 1. Samples in the same class are not connected.

The elements of an adjacency matrix \mathbf{W} can be defined as follows:

$$w_{ij} = \begin{cases} 1, & y_i \neq y_j, \\ 0, & y_i = y_j. \end{cases} \quad (4)$$

Here, $y_i \in \mathbf{Y}$ is the label of \mathbf{x}_i .

According to the above definition of the adjacency matrix \mathbf{W} , we can see that the distance of the samples in the same class can be considered 0. For samples in different classes, the distance between them can be considered 1.

Then, the relationship graph matrix can be expressed as

$$\mathbf{L} = \mathbf{D} - \mathbf{W}, \quad (5)$$

where $\mathbf{D} = \text{diag}(\mathbf{W} \cdot \mathbf{1}_{n \times 1})$ is the degree matrix; $\mathbf{1}_{n \times 1}$ stands for a vector with $n \times 1$, whose elements are exactly 1; n is the number of training samples.

As for the imbalanced dataset \mathbf{X} , we need to increase the loss of misclassification of positive samples because the misclassification of positive samples (patients with epilepsy) could cause serious consequences. This can be realized by regulating the degree matrix \mathbf{D} . The shortcomings of the cost learning algorithm can be compensated by increasing the relationship between samples. Therefore, the relationship graph not only ensures the accuracy of positive sample classification but also makes up for the lack of the mutual relationships and prior knowledge between samples.

According to the above description, the relationship graph matrix \mathbf{L} of the synthetic dataset in Figure 4 can be expressed as

$$\mathbf{L} = 4 \begin{bmatrix} & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & & 2 & 0 & 0 & 0 & 0 & -1 & -1 \\ 2 & & 0 & 2 & 0 & 0 & 0 & -1 & -1 \\ 3 & & 0 & 0 & 2 & 0 & 0 & -1 & -1 \\ 4 & & 0 & 0 & 0 & 2 & 0 & -1 & -1 \\ 5 & & 0 & 0 & 0 & 0 & 2 & -1 & -1 \\ 6 & & -1 & -1 & -1 & -1 & -1 & 5 & 0 \\ 7 & & -1 & -1 & -1 & -1 & -1 & 0 & 5 \end{bmatrix} \quad (6)$$

3.2. Objective Function of G-ELM. According to the above relationship graph and ELM, the objective function of the G-ELM can be expressed as follows:

$$\begin{aligned} \min \quad & \frac{1}{2} \|\boldsymbol{\beta}\|^2 + \frac{1}{2} C \sum_{i=1}^n \sum_{j=1}^n \varepsilon_i L_{ij} \varepsilon_j, \\ \text{s.t.} \quad & \mathbf{h}(\mathbf{x}_i) \boldsymbol{\beta} = y_i - \varepsilon_i, \quad i = 1, 2, \dots, n. \end{aligned} \quad (7)$$

Here, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$, n is the number of samples in \mathbf{X} , d is the sample dimension, and $\mathbf{Y} = [y_1, y_2, \dots, y_n]^T$ represents the true class label of the samples. \mathbf{H} and $\mathbf{h}(\mathbf{x}_i)$ are the same as defined in ELM. $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_m]^T$ represents the output weight vector. $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n]^T$ represents the loss between the network outputs and the target outputs.

$$\mathbf{L} = \begin{bmatrix} L_{11} & \cdots & L_{1n} \\ \vdots & \ddots & \vdots \\ L_{n1} & \cdots & L_{nn} \end{bmatrix}_{n \times n}. \quad (8)$$

Equation (8) is the relationship graph matrix of the samples.

By comparing (7) with (1), we can see that G-ELM is an improved version of ELM and still has the characteris-

tics of high learning speed and strong generalization ability from ELM.

3.3. Solution of G-ELM. In this subsection, we attempt to optimize the objective function of G-ELM. According to [20], the objective function of G-ELM is a convex optimization problem. The specific optimization solution process is as follows:

The Lagrangian function corresponding to (7) is

$$J = \frac{1}{2} \|\boldsymbol{\beta}\|^2 + \frac{1}{2} C \sum_{i=1}^n \sum_{j=1}^n \varepsilon_i L_{ij} \varepsilon_j - \sum_{i=1}^n \alpha_i (\mathbf{h}(\mathbf{x}_i) \boldsymbol{\beta} - y_i + \varepsilon_i). \quad (9)$$

Let the derivation of J with respect to $\boldsymbol{\beta}$, ε_i , α_i equal to zero:

$$\frac{\partial J}{\partial \boldsymbol{\beta}} = \boldsymbol{\beta} - \sum_{i=1}^n \alpha_i \mathbf{h}(\mathbf{x}_i) = 0 \Rightarrow \boldsymbol{\beta} = \mathbf{H}^T \boldsymbol{\alpha}, \quad (10a)$$

$$\frac{\partial J}{\partial \varepsilon_i} = -\alpha_i + C \sum_{j=1}^n L_{ij} \varepsilon_j = 0 \Rightarrow \boldsymbol{\alpha} = \mathbf{C} \mathbf{L} \boldsymbol{\varepsilon}, \quad (10b)$$

$$\frac{\partial J}{\partial \alpha_i} = \mathbf{h}(\mathbf{x}_i) \boldsymbol{\beta} - y_i + \varepsilon_i = 0 \Rightarrow \mathbf{H} \boldsymbol{\beta} - \mathbf{Y} + \boldsymbol{\varepsilon} = 0. \quad (10c)$$

Substituting (10a) and (10b) into (10c), we obtain

$$\begin{aligned} \mathbf{H} \mathbf{H}^T \boldsymbol{\alpha} - \mathbf{Y} + \frac{1}{C} \mathbf{L}^{-1} \boldsymbol{\alpha} &= 0 \Rightarrow \left(\mathbf{H} \mathbf{H}^T + \frac{1}{C} \mathbf{L}^{-1} \right) \boldsymbol{\alpha} = \mathbf{Y} \\ &\Rightarrow \left(\mathbf{L} \mathbf{H} \mathbf{H}^T + \frac{\mathbf{I}}{C} \right) \boldsymbol{\alpha} = \mathbf{L} \mathbf{Y}. \end{aligned} \quad (11)$$

Combining (10a) and (11), we obtain

$$\boldsymbol{\beta} = \begin{cases} \mathbf{H}^T \left(\mathbf{L} \mathbf{H} \mathbf{H}^T + \frac{\mathbf{I}}{C} \right)^{-1} \mathbf{L} \mathbf{Y} & n < m, \\ \left(\mathbf{H}^T \mathbf{L} \mathbf{H} + \frac{\mathbf{I}}{C} \right)^{-1} \mathbf{H}^T \mathbf{L} \mathbf{Y} & n \geq m. \end{cases} \quad (12)$$

With the obtained solution, i.e., $\boldsymbol{\beta}^*$, the predicted class label of the testing sample can be obtained as follows:

$$\mathbf{y}_{\text{test}} = \text{sign}(\mathbf{x}_{\text{test}} \boldsymbol{\beta}^*), \quad (13)$$

where \mathbf{x}_{test} is a testing sample.

3.4. Learning Algorithm of G-ELM. According to the above derivation, the implementation of G-ELM is summarized in Algorithm 1.

4. Experiments

To demonstrate the effectiveness of the proposed G-ELM, we conducted extensive experiments on a real epileptic EEG dataset. The proposed G-ELM was verified by comparing it with five ELM-based methods, i.e., ELM [19], W1-ELM [20], W2-ELM [20], R1-ELM [25], and R2-ELM [25], using

Input: The training samples $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$ and their corresponding labels $\mathbf{Y} = [y_1, y_2, \dots, y_n]^T$, where $\mathbf{x}_i \in \mathbb{R}^d (i = 1, 2, \dots, n)$. The number of the hidden nodes m ; the input weights $\mathbf{A} \in \mathbb{R}^{m \times d}$ and input biases $\mathbf{b} \in \mathbb{R}^m$; the penalty parameter C .

Output: The predicted class label of the testing sample \mathbf{x}_{test} .

Step 1: Construct the mapping matrix of hidden layer \mathbf{H} according to Eq. (2).

Step 2: Compute the relationship graph matrix corresponding to the training samples \mathbf{X} according to Eq. (4) and Eq. (5).

Step 3:
If $n < m$
Then compute the output weight vector $\boldsymbol{\beta}^*$ using the first formula in Eq. (12).
Else compute $\boldsymbol{\beta}^*$ using the second formula in Eq. (12).

Step 4: Return the predicted class label of the testing sample $\mathbf{y}_{test} = \text{sign}(\mathbf{x}_{test} \boldsymbol{\beta}^*)$.

ALGORITHM 1: G-ELM.

TABLE 2: Nine datasets constructed from *Bonn* for the experiments.

Datasets	Negative classes	Positive class	Sizes	IR
D1	A, C	E	300	2
D2	B, C	E	300	2
D3	A, D	E	300	2
D4	B, D	E	300	2
D5	A, B, C	E	400	3
D6	A, B, D	E	400	3
D7	A, C, D	E	400	3
D8	B, C, D	E	400	3
D9	A, B, C, D	E	500	4

six imbalanced classification evaluation indices and average standard deviation on the real *Bonn* dataset. Except for ELM, the other comparison methods are imbalanced classification methods. All the experiments were conducted on a computer with Intel Core i5-3317U 1.70 GHz CPU and 16 GB RAM by using MATLAB 2016a. The details of the experimental settings and results are presented in the following sections.

4.1. Data Preparation. Although the real *Bonn* dataset has been used in many studies, the way of using it in this study differs from those in previous works. To evaluate the performance of the proposed G-ELM, nine imbalanced datasets were generated from the original five groups of EEG signals to simulate the imbalanced classification scenario. The details of the nine datasets are summarized in Table 2. In each dataset, the EEG signals of patients with epilepsy (E) were regarded as a positive class, while the other groups were regarded as a negative class, to identify whether the patients with epilepsy are experiencing seizure activity. A brief description of the five groups (A, B, C, D, and E) can be found in Table 1. The last column of Table 2 is IR, which is used to show the degree of imbalance of the dataset. IR can be defined as follows:

$$\text{IR} = \frac{n^-}{n^+}, \quad (14)$$

where n_+ and n_- represent the number of samples of the positive class and the negative class, respectively.

In our experiment, we randomly partitioned each dataset. In each dataset, 80% of the dataset were used for training and the remaining 20% were used for testing.

4.2. Evaluation Indices. In our experiments, we used six imbalanced classification evaluation indices to evaluate all the adopted methods. The six imbalanced classification evaluation indices were accuracy, precision, recall, *F*-measure, *G*_means, and AUC, which can be, respectively, defined as

$$\text{Accuracy} = \frac{(\text{TN} + \text{TP})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})},$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})},$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})},$$

$$F\text{-measure} = \frac{(2 \cdot \text{recall} \cdot \text{precision})}{(\text{recall} + \text{precision})},$$

$$G\text{-means} = \sqrt{\frac{\text{TP}}{(\text{TP} + \text{FN})} \cdot \frac{\text{TN}}{(\text{TN} + \text{FP})}}. \quad (15)$$

Here, TP is the number of true positive samples, FN is the number of false negative samples, TN is the number of true negative samples, and FP is the number of false positive samples, respectively.

$$\text{AUC} = \sum_{i \in \mathcal{N}_+} \sum_{j \in \mathcal{N}_-} \frac{I(P(\mathbf{x}_i), P(\mathbf{x}_j))}{n_+ n_-}, \quad (16)$$

where \mathcal{N}_+ is the set of all the indexes of the positive samples and \mathcal{N}_- is the set of those of the negative samples; $n_+ = |\mathcal{N}_+|$ and $n_- = |\mathcal{N}_-|$. $P(\mathbf{x})$ is the prediction value of \mathbf{x} . $I(\cdot)$ is the indicator function

$$I(P(\mathbf{x}_i), P(\mathbf{x}_j)) = \begin{cases} 1, & P(\mathbf{x}_i) > P(\mathbf{x}_j), \\ 0.5, & P(\mathbf{x}_i) = P(\mathbf{x}_j), \\ 0, & P(\mathbf{x}_i) < P(\mathbf{x}_j). \end{cases} \quad (17)$$

TABLE 3: Results of all methods on the accuracy index.

Datasets	ELM	W1-ELM	W2-ELM	R1-ELM	R2-ELM	G-ELM
D1	0.8903(0.0027)	0.8593(0.0149)	0.8815(0.0176)	0.9389(0.0012)	0.9178(0.0041)	0.9489(0.0011)
D2	0.8744(0.0033)	0.8513(0.0150)	0.8574(0.0139)	0.9333(0.0021)	0.9207(0.0025)	0.9533(0.0023)
D3	0.8867(0.0056)	0.8472(0.0108)	0.8513(0.0125)	0.9067(0.0028)	0.8989(0.0022)	0.9511(0.0011)
D4	0.8633(0.0052)	0.8593(0.0153)	0.8452(0.0179)	0.9167(0.0019)	0.8989(0.0024)	0.9433(0.0024)
D5	0.7475(0.0081)	0.8594(0.0144)	0.8674(0.0229)	0.8917(0.0034)	0.7625(0.0058)	0.9383(0.0035)
D6	0.8325(0.0097)	0.8712(0.0167)	0.8657(0.0201)	0.9217(0.0031)	0.8258(0.0109)	0.9425(0.0028)
D7	0.8783(0.0035)	0.8774(0.0192)	0.8681(0.0088)	0.9283(0.0020)	0.9158(0.0023)	0.9550(0.0018)
D8	0.9042(0.0045)	0.8519(0.0166)	0.8625(0.0128)	0.9225(0.0036)	0.9125(0.0036)	0.9533(0.0018)
D9	0.9102(0.0041)	0.8675(0.0151)	0.8658(0.0127)	0.9180(0.0039)	0.8880(0.0029)	0.9480(0.0033)
Ave	0.8653	0.8605	0.8628	0.9198	0.8823	0.9482
Ave. std	0.0052	0.0153	0.0155	0.0027	0.0041	0.0022

TABLE 4: Results of all methods on the recall index.

Datasets	ELM	W1-ELM	W2-ELM	R1-ELM	R2-ELM	G-ELM
D1	0.7467(0.0072)	0.8231(0.0194)	0.8048(0.0219)	0.8533(0.0056)	0.8967(0.0038)	0.9367(0.0021)
D2	0.7233(0.0083)	0.8051(0.0351)	0.7881(0.0194)	0.8467(0.0034)	0.8667(0.0031)	0.9033(0.0038)
D3	0.8403(0.0088)	0.7933(0.0172)	0.7947(0.0184)	0.9233(0.0045)	0.9233(0.0055)	0.9533(0.0034)
D4	0.7867(0.0064)	0.7625(0.0199)	0.8156(0.0354)	0.8767(0.0045)	0.8833(0.0047)	0.9196(0.0034)
D5	0.7167(0.0110)	0.8006(0.0201)	0.8509(0.0218)	0.7967(0.0080)	0.8567(0.0045)	0.8489(0.0035)
D6	0.6867(0.0134)	0.8037(0.0270)	0.8374(0.0166)	0.8333(0.0054)	0.8498(0.0065)	0.8805(0.0056)
D7	0.7405(0.0093)	0.7825(0.0273)	0.7751(0.0162)	0.8597(0.0042)	0.8906(0.0045)	0.8933(0.0028)
D8	0.7667(0.0054)	0.7619(0.0254)	0.8125(0.0272)	0.8933(0.0053)	0.9033(0.0058)	0.9267(0.0028)
D9	0.6902(0.0114)	0.7972(0.0150)	0.8143(0.0172)	0.8533(0.0034)	0.8667(0.0054)	0.8896(0.0045)
Ave	0.7442	0.7922	0.8104	0.8596	0.8819	0.9058
Ave. std	0.0090	0.0229	0.0216	0.0049	0.0049	0.0035

TABLE 5: Results of all methods on the precision index.

Datasets	ELM	W1-ELM	W2-ELM	R1-ELM	R2-ELM	G-ELM
D1	0.9078(0.0058)	0.8247(0.0304)	0.8342(0.0301)	0.9619(0.0087)	0.8648(0.0097)	0.9126(0.0029)
D2	0.8801(0.0077)	0.7878(0.0313)	0.7901(0.0298)	0.9497(0.0077)	0.8914(0.0063)	0.9569(0.0091)
D3	0.8279(0.0125)	0.8009(0.0261)	0.7911(0.0212)	0.8208(0.0062)	0.8040(0.0053)	0.9053(0.0026)
D4	0.8054(0.0124)	0.8725(0.0375)	0.7729(0.0290)	0.8761(0.0079)	0.8267(0.0060)	0.9131(0.0085)
D5	0.5031(0.0107)	0.8438(0.0360)	0.8609(0.0434)	0.7845(0.0147)	0.5178(0.0078)	0.9007(0.0102)
D6	0.6825(0.0243)	0.8833(0.0388)	0.8502(0.0399)	0.8541(0.0112)	0.6244(0.0216)	0.8952(0.0125)
D7	0.7714(0.0131)	0.7905(0.0400)	0.8817(0.0349)	0.8568(0.0088)	0.7985(0.0067)	0.9266(0.0092)
D8	0.8431(0.0149)	0.8466(0.0359)	0.7732(0.0233)	0.8194(0.0120)	0.7860(0.0119)	0.8932(0.0083)
D9	0.8374(0.0176)	0.8414(0.0204)	0.8229(0.0367)	0.7701(0.0139)	0.6733(0.0087)	0.8650(0.0174)
Ave	0.7843	0.8324	0.8197	0.8548	0.7541	0.9076
Ave. std	0.0132	0.0329	0.0320	0.0101	0.0093	0.0090

F -measure, G _means, and AUC are commonly used imbalanced classification evaluation indices [42–45].

4.3. Adopted Methods and Parameter Settings. In the experiments, five ELM-based methods, i.e., ELM [19], W1-ELM [20], W2-ELM [20], R1-ELM [25], and R2-ELM [25], were adopted for comparisons with G-ELM. Referring to the guidelines in [2, 20, 46], a grid search strategy based

on G _means was used to determine appropriate parameters of all the methods. We set parameter C in the range of $2^{[-28:28]}$ and parameter m in the range of $\{50, 100, 300, 500, 1000\}$ for all the adopted methods. All the adopted methods were run ten times on each generated imbalanced dataset. The average experimental results corresponding to the six imbalanced classification evaluation indices are reported.

TABLE 6: Results of all methods on the F -measure index.

Datasets	ELM	W1-ELM	W2-ELM	R1-ELM	R2-ELM	G-ELM
D1	0.8187(0.0047)	0.7716(0.0111)	0.8103(0.0218)	0.9031(0.0013)	0.8797(0.0053)	0.9244(0.0016)
D2	0.7931(0.0057)	0.7561(0.0225)	0.7808(0.0151)	0.8946(0.0028)	0.8786(0.0035)	0.9285(0.0031)
D3	0.8323(0.0075)	0.7620(0.0130)	0.7680(0.0151)	0.8686(0.0036)	0.8590(0.0028)	0.9285(0.0017)
D4	0.7942(0.0062)	0.7628(0.0152)	0.7539(0.0220)	0.8755(0.0021)	0.8536(0.0033)	0.9158(0.0032)
D5	0.5883(0.0073)	0.7463(0.0123)	0.7636(0.0188)	0.7872(0.0044)	0.6444(0.0054)	0.8739(0.0053)
D6	0.6762(0.0121)	0.8032(0.0250)	0.8285(0.0231)	0.8423(0.0052)	0.7148(0.0134)	0.8854(0.0040)
D7	0.7528(0.0061)	0.7924(0.0261)	0.7507(0.0116)	0.8575(0.0033)	0.8412(0.0037)	0.9089(0.0032)
D8	0.8012(0.0074)	0.7708(0.0233)	0.7797(0.0176)	0.8531(0.0056)	0.8387(0.0053)	0.9089(0.0030)
D9	0.7545(0.0108)	0.7130(0.0178)	0.7069(0.0149)	0.8079(0.0074)	0.7566(0.0044)	0.8744(0.0068)
Ave	0.7568	0.7642	0.7714	0.8544	0.8074	0.9054
Ave. std	0.0075	0.0185	0.0178	0.0040	0.0052	0.0036

TABLE 7: Results of all methods on the G _means index.

Datasets	ELM	W1-ELM	W2-ELM	R1-ELM	R2-ELM	G-ELM
D1	0.8471(0.0040)	0.8242(0.0102)	0.8516(0.0174)	0.9150(0.0012)	0.9122(0.0033)	0.9458(0.0012)
D2	0.8286(0.0047)	0.8318(0.0197)	0.8402(0.0125)	0.9092(0.0016)	0.9057(0.0023)	0.9399(0.0014)
D3	0.8738(0.0055)	0.8213(0.0105)	0.8253(0.0120)	0.9106(0.0025)	0.9046(0.0022)	0.9516(0.0015)
D4	0.8417(0.0041)	0.8010(0.0108)	0.8201(0.0182)	0.9059(0.0011)	0.8948(0.0024)	0.9372(0.0016)
D5	0.7352(0.0055)	0.8270(0.0081)	0.8188(0.0147)	0.8570(0.0022)	0.7909(0.0038)	0.9069(0.0025)
D6	0.7757(0.0074)	0.8496(0.0169)	0.8346(0.0147)	0.8901(0.0028)	0.8326(0.0073)	0.9203(0.0004)
D7	0.8265(0.0046)	0.8518(0.0179)	0.8151(0.0073)	0.9043(0.0016)	0.9069(0.0022)	0.9334(0.0009)
D8	0.8532(0.0037)	0.8203(0.0178)	0.8459(0.0134)	0.9123(0.0025)	0.9091(0.0023)	0.9442(0.0007)
D9	0.8154(0.0071)	0.8452(0.0124)	0.8048(0.0069)	0.8927(0.0025)	0.8796(0.0020)	0.9253(0.0009)
Ave	0.8219	0.8303	0.8285	0.8997	0.8818	0.9338
Ave. std	0.0052	0.0138	0.0130	0.0020	0.0031	0.0012

TABLE 8: Results of all methods on the AUC index.

Datasets	ELM	W1-ELM	W2-ELM	R1-ELM	R2-ELM	G-ELM
D1	0.8542(0.0035)	0.8371(0.0088)	0.8679(0.0167)	0.9175(0.0009)	0.9125(0.0034)	0.9458(0.0012)
D2	0.8367(0.0042)	0.8418(0.0159)	0.8438(0.0117)	0.9117(0.0016)	0.9067(0.0023)	0.9408(0.0015)
D3	0.8750(0.0055)	0.8409(0.0102)	0.8319(0.0120)	0.9108(0.0025)	0.9050(0.0022)	0.9517(0.0015)
D4	0.8442(0.0042)	0.8162(0.0113)	0.8546(0.0176)	0.9067(0.0011)	0.8950(0.0024)	0.9375(0.0016)
D5	0.7372(0.0054)	0.8296(0.0070)	0.8239(0.0128)	0.8607(0.0020)	0.7939(0.0035)	0.9089(0.0025)
D6	0.7839(0.0071)	0.8616(0.0156)	0.8403(0.0147)	0.8922(0.0027)	0.8339(0.0072)	0.9217(0.0004)
D7	0.8322(0.0041)	0.8569(0.0166)	0.8165(0.0067)	0.9056(0.0016)	0.9072(0.0022)	0.9344(0.0009)
D8	0.8583(0.0037)	0.8319(0.0167)	0.8561(0.0114)	0.9128(0.0025)	0.9094(0.0023)	0.9444(0.0007)
D9	0.8275(0.0062)	0.8494(0.0119)	0.7990(0.0070)	0.8938(0.0026)	0.8794(0.0020)	0.9263(0.0010)
Ave	0.8277	0.8406	0.8371	0.9013	0.8826	0.9346
Ave. std	0.0049	0.0127	0.0123	0.0019	0.0031	0.0013

4.4. *Experimental Results.* To evaluate the classification performance of the proposed G-ELM, five ELM-based methods were used for performance comparison. All experiments were repeated ten times for fairness. The mean and standard deviation of the corresponding indices of all methods in each dataset are reported in Tables 3–8. The best results are shown in bold. The improvement of G-ELM relative to ELM on all

datasets using the six imbalanced classification evaluation indices is shown in Figure 5.

According to experimental results in Tables 3–8, the following observations can be made:

- (1) For the adopted six imbalanced classification evaluation indices, the proposed G-ELM performs best on

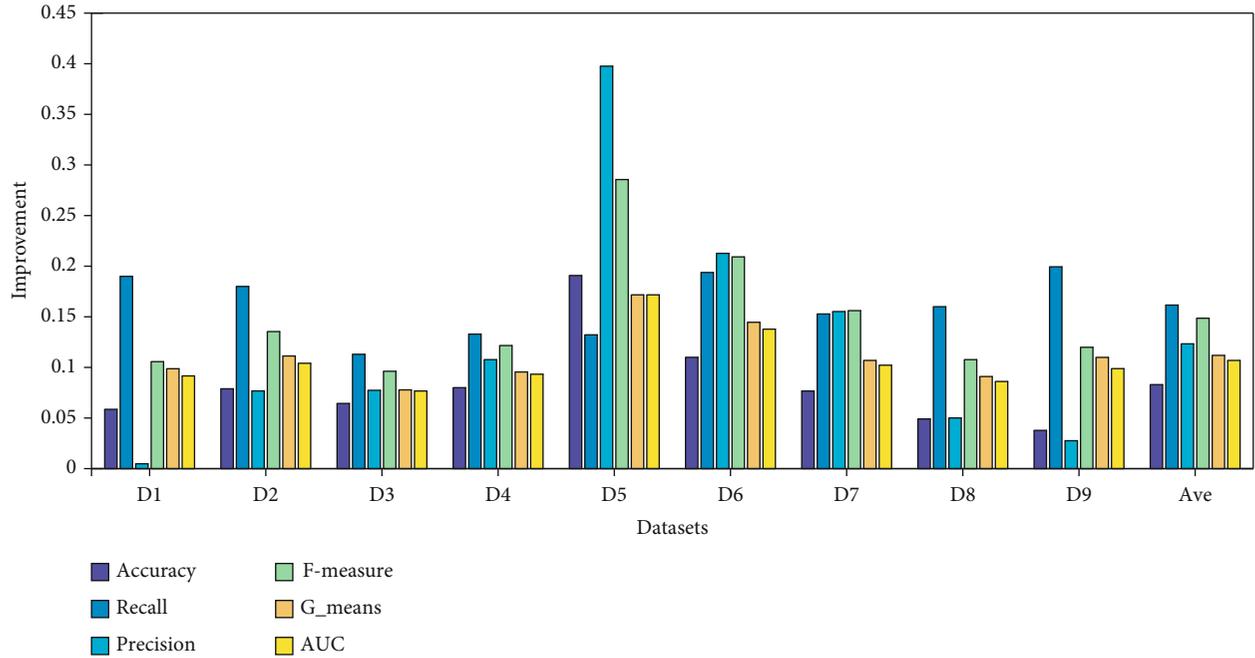


FIGURE 5: Improvement of G-ELM compared with ELM across six indices on all datasets.

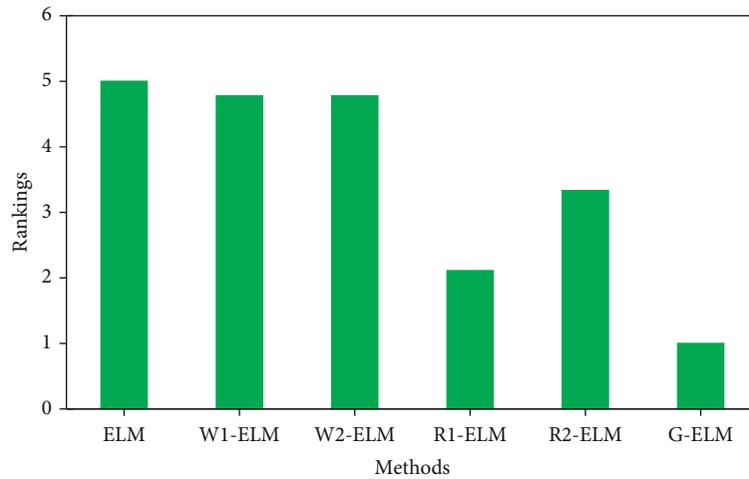


FIGURE 6: Rankings of all methods (Friedman test).

most datasets. This is because G-ELM can suppress the misclassification of negative samples while considering the accuracy of positive samples and has a high classification performance, which is suitable for imbalanced epileptic EEG signal recognition

- (2) In general, G-ELM, R1-ELM, R2-ELM, W1-ELM, and W2-ELM achieved better performances than ELM. This is due to the addition of a cost matrix, which makes them more suitable for imbalanced classification. Moreover, G-ELM has the best effect because it adds sample information using a relationship graph
- (3) Tables 4 and 5 show the results of recall and precision of all methods. They evaluate the performance of

TABLE 9: The post hoc hypothesis test results ($\alpha_{\text{Fri}} = 0.05$).

i	Methods	z	p_{Fri}	Holm	Hypothesis
5	ELM	4.535574	0.000006	0.01	Rejected
4	W1-ELM	4.283597	0.000018	0.0125	Rejected
3	W2-ELM	4.283597	0.000018	0.016667	Rejected
2	R2-ELM	2.645751	0.008151	0.025	Rejected
1	R1-ELM	1.259882	0.207712	0.05	Not rejected

imbalanced classification from two different perspectives. From the excellent performance of G-ELM in Tables 4 and 5, we can see that adding information regarding relationships between samples can increase the weight of positive samples

- (4) F -measure and G _means in Tables 6 and 7 are two important indices to measure the performance of imbalanced classification methods, which can be combined with recall and precision to evaluate the effect of the methods. From the results, we can see that the proposed G-ELM has the best performance. It has excellent performance in imbalanced epileptic EEG signal recognition
- (5) AUC is an important index to evaluate imbalanced classifiers. From Table 8, we can see that the performance of G-ELM on all datasets is the best. G-ELM has excellent performance in imbalanced classification and good effectiveness of imbalanced epileptic EEG signal recognition

4.5. Statistical Analysis. Statistical analysis was performed to further analyze the performances of all the adopted methods in our experiments. For conciseness, we only present statistical analysis of the G _means results. Firstly, the Friedman test [47] was used to calculate the average ranking of each method. The rankings of all the adopted methods are shown in Figure 6. In Figure 6, we can see that the performance of G-ELM is the best.

Then, the post hoc hypothesis test [48] was used to evaluate the statistical significance of the performance differences between G-ELM and the other adopted methods. Post hoc hypothesis test results ($\alpha_{\text{Fri}} = 0.05$) are presented in Table 9. In Table 9, we can see that the null hypothesis is rejected when $p_{\text{Fri}} \leq 0.025$ due to $p_{\text{Fri}} \leq \text{Holm}$. Therefore, performance differences between G-ELM and the other adopted methods are significant, which means that G-ELM is effective for imbalanced epileptic EEG signal recognition.

5. Conclusions

In this study, we aimed to address the challenge that traditional machine learning methods ignore the imbalance of epileptic EEG datasets and the existing imbalanced classification methods ignore the relationships between samples. A graph-based ELM was proposed for imbalanced epileptic EEG signal recognition. First, graph theory was used to construct the relationship between samples according to the distribution. Second, a model combining the relationship graph and ELM was proposed; this model inherited the rapid learning and good generalization capabilities of ELM while maintaining satisfactory classification. Experiments on a real imbalanced epileptic EEG dataset demonstrated the effectiveness and applicability of the proposed method. However, there is still room for improvement in the scope and search method of the optimal parameters in this experiment. In the future, ways to design a better method to determine the optimal parameters will be further studied and explored.

Data Availability

Data can be downloaded from <http://www.epileptologie-bonn.de/cms/upload/workgroup/lehnerz/eegdata.html>.

Conflicts of Interest

None of the authors have any conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61772198 and by the Natural Science Foundation of Jiangsu Province under Grant BK20161268.

References

- [1] S. Benbadis and W. Hauser, "An estimate of the prevalence of psychogenic non-epileptic seizures," *Seizure*, vol. 9, no. 4, pp. 280–281, 2000.
- [2] C. Yang, Z. Deng, K. Choi, and S. Wang, "Takagi–Sugeno–Kang transfer learning fuzzy logic system for the adaptive recognition of epileptic electroencephalogram signals," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 5, pp. 1079–1094, 2016.
- [3] Z. Jiang, F.-L. Chung, and S. Wang, "Recognition of multi-class epileptic EEG signals based on knowledge and label space inductive transfer," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 4, pp. 630–642, 2019.
- [4] M. Z. Parvez and M. Paul, "Epileptic seizure prediction by exploiting spatiotemporal relationship of EEG signals using phase correlation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 158–168, 2016.
- [5] Z. Iscan, Z. Dokur, and T. Demiralp, "Classification of electroencephalogram signals with combined time and frequency features," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10499–10505, 2011.
- [6] T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in eegs using time-frequency analysis," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 703–710, 2009.
- [7] Z. Yang, K. C. Deng, K.-S. Choi, Y. Jiang, and S. Wang, "Transductive domain adaptive learning for epileptic electroencephalogram recognition," *Artificial Intelligence in Medicine*, vol. 62, no. 3, pp. 165–177, 2014.
- [8] Y. Zhang, J. Dong, J. Zhu, and C. Wu, "Common and special knowledge-driven TSK fuzzy system and its modeling and application for epileptic EEG signals recognition," *IEEE Access*, vol. 7, pp. 127600–127614, 2019.
- [9] F. P. Lestari, M. Haekal, R. E. Edison, F. R. Fauzy, S. N. Khotimah, and F. Haryanto, "Epileptic Seizure Detection in EEGs by Using Random Tree Forest, Naïve Bayes and KNN Classification," *Journal of Physics: Conference Series*, vol. 1505, no. 1, p. 12055, 2020.
- [10] H. Choubey and A. Pandey, "A combination of statistical parameters for the detection of epilepsy and EEG classification using ANN and KNN classifier," *Signal, Image and Video Processing*, vol. 15, no. 3, pp. 475–483, 2021.
- [11] Y. Zhang, Y. Zhang, J. Wang, and X. Zheng, "Comparison of classification methods on EEG signals based on wavelet packet decomposition," *Neural Computing and Applications*, vol. 26, no. 5, pp. 1217–1225, 2015.
- [12] Y. Zhang, H. Ishibuchi, and S. Wang, "Deep Takagi–Sugeno–Kang fuzzy classifier with shared linguistic fuzzy rules," *IEEE*

- Transactions on Fuzzy Systems*, vol. 26, no. 3, pp. 1535–1549, 2018.
- [13] Z. Deng, P. Xu, L. Xie, K.-S. Choi, and S. Wang, “Transductive joint-knowledge-transfer TSK FS for recognition of epileptic EEG signals,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 8, pp. 1481–1494, 2018.
- [14] Q. Yuan, W. Zhou, S. Li, and D. Cai, “Epileptic EEG classification based on extreme learning machine and nonlinear features,” *Epilepsy Research*, vol. 96, no. 1-2, pp. 29–38, 2011.
- [15] L. Shi and B. Lu, “EEG-based vigilance estimation using extreme learning machines,” *Neurocomputing*, vol. 102, no. 15, pp. 135–143, 2013.
- [16] Q. Yuan, W. Zhou, L. Zhang et al., “Epileptic seizure detection based on imbalanced classification and wavelet packet transform,” *Seizure*, vol. 50, pp. 99–108, 2017.
- [17] M. K. Siddiqui, X. Huang, R. Morales-Menendez, N. Hussain, and K. Khatoun, “Machine learning based novel cost-sensitive seizure detection classifier for imbalanced EEG data sets,” *International Journal on Interactive Design and Manufacturing*, vol. 14, no. 4, pp. 1491–1509, 2020.
- [18] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: a new learning scheme of feedforward neural networks,” in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, vol. 2, pp. 985–990, Budapest, Hungary, 2004.
- [19] G.-B. Huang, Q. Y. Zhu, and C. K. Siew, “Extreme learning machine: theory and applications,” *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.
- [20] W. Zong, G.-B. Huang, and Y. Chen, “Weighted extreme learning machine for imbalance learning,” *Neurocomputing*, vol. 101, pp. 229–242, 2013.
- [21] W. B. Zhang and H. B. Ji, “Fuzzy extreme learning machine for classification,” *Electronics Letters*, vol. 49, no. 7, pp. 448–450, 2013.
- [22] H. Yu, C. Sun, X. Yang, W. Yang, J. Shen, and Y. Qi, “ODOC-ELM: optimal decision outputs compensation-based extreme learning machine for classifying imbalanced data,” *Knowledge-Based Systems*, vol. 92, pp. 55–70, 2016.
- [23] K. Li, X. Kong, Z. Lu, L. Wenyin, and J. Yin, “Boosting weighted elm for imbalanced learning,” *Neurocomputing*, vol. 128, pp. 15–21, 2014.
- [24] J. Yang, H. Yu, X. Yang, and X. Zuo, “Imbalanced extreme learning machine based on probability density estimation,” in *International Workshop on Multi-disciplinary Trends in Artificial Intelligence*, pp. 160–167, Springer, Cham, 2015.
- [25] S. Shukla and R. N. Yadav, “Regularized weighted circular complex-valued extreme learning machine for imbalanced learning,” *IEEE Access*, vol. 3, pp. 3048–3057, 2015.
- [26] W. Xiao, J. Zhang, Y. Li, and W. Yang, “Imbalanced extreme learning machine for classification with imbalanced data distributions,” *Proceedings in Adaptation, Learning and Optimization*, vol. 7, pp. 503–514, 2016.
- [27] J. Du, C. Vong, Y. Chang, and Y. Jiao, “Online sequential extreme learning machine with under-sampling and over-sampling for imbalanced big data classification,” *Proceedings of ELM-2016*, 2018.
- [28] Y. Zhang, B. Liu, J. Cai, and S. Zhang, “Ensemble weighted extreme learning machine for imbalanced data classification based on differential evolution,” *Neural Computing and Applications*, vol. 28, Supplement 1, pp. 259–267, 2017.
- [29] W. Xiao, J. Zhang, Y. Li, S. Zhang, and W. Yang, “Class-specific cost regulation extreme learning machine for imbalanced classification,” *Neurocomputing*, vol. 261, pp. 70–82, 2017.
- [30] H. Yu, C. Sun, X. Yang, S. Zheng, Q. Wang, and X. Xi, “LW-ELM: a fast and flexible cost-sensitive learning framework for classifying imbalanced data,” *IEEE Access*, vol. 6, pp. 28488–28500, 2018.
- [31] B. S. Raghuwanshi and S. Shukla, “Class-specific extreme learning machine for handling binary class imbalance problem,” *Neural Networks*, vol. 105, pp. 206–217, 2018.
- [32] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, “Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state,” *Physical Review E*, vol. 64, no. 6, article 061907, 2001.
- [33] T. Strutz, *Data Fitting and Uncertainty: A Practical Introduction to Weighted Least Squares and Beyond*, Vieweg, Wiesbaden, Germany, 2010.
- [34] G. Li, K. Chang, and S. C. H. Hoi, “Multiview semi-supervised learning with consensus,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 11, pp. 2040–2051, 2012.
- [35] F. Zhu, L. Shao, and M. Lin, “Multi-view action recognition using local similarity random forests and sensor fusion,” *Pattern Recognition Letters*, vol. 34, no. 1, pp. 20–24, 2013.
- [36] B. R. Greene, S. Faul, W. P. Marnane, G. Lightbody, I. Korotchikova, and G. B. Boylan, “A comparison of quantitative EEG features for neonatal seizure detection,” *Clinical Neurophysiology*, vol. 119, no. 6, pp. 1248–1261, 2008.
- [37] E. A. Vivaldi and A. Bassi, “Frequency domain analysis of sleep eeg for visualization and automated state detection,” in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3740–3743, New York, NY, USA, August 2006.
- [38] C. W. N. F. C. W. Fadzal, W. Mansor, L. Y. Khuan, and A. Zabidi, “Short-time Fourier transform analysis of EEG signal from writing,” in *2012 IEEE 8th International Colloquium on Signal Processing and its Applications*, pp. 525–527, Malacca, Malaysia, March 2012.
- [39] H. Ocak, “Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy,” *Expert Systems with Applications*, vol. 36, no. 2, pp. 2027–2036, 2009.
- [40] D. Hu, W. Li, and X. Chen, “Feature extraction of motor imagery EEG signals based on wavelet packet decomposition,” in *The 2011 IEEE/ICME International Conference on Complex Medical Engineering*, pp. 694–697, Harbin, China, May 2011.
- [41] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken, “Multilayer feedforward networks with a nonpolynomial activation function can approximate any function,” *Neural Networks*, vol. 6, no. 6, pp. 861–867, 1993.
- [42] Z. Zhu, Z. Wang, D. Li, and W. Du, “Multiple empirical kernel learning with majority projection for imbalanced problems,” *Applied Soft Computing*, vol. 76, pp. 221–236, 2019.
- [43] O. M. Younis, M. M. Krunz, and S. Ramasubramanian, “ROC: resilient online coverage for surveillance applications,” *IEEE/ACM Transactions on Networking*, vol. 19, no. 1, pp. 251–264, 2011.

- [44] Y. Zhang, S. Wang, K. Xia, Y. Jiang, and P. Qian, "Alzheimer's disease multiclass diagnosis via multimodal neuroimaging embedding feature selection and fusion," *Information Fusion*, vol. 66, pp. 170–183, 2021.
- [45] T. A. Joachims, "A support vector method for multivariate performance measures," in *Proceedings of the 22nd International Conference on Machine Learning - ICML '05*, pp. 377–384, Bonn, Germany, 2005.
- [46] J. Zhou, Z. Jiang, and S. Wang, "Laplacian least learning machine with dynamic updating for imbalanced classification," *Applied Soft Computing*, vol. 88, p. 106028, 2020.
- [47] Y. Zhang, F.-L. Chung, and S. Wang, "Takagi-sugeno-kang fuzzy systems with dynamic rule weights," *Journal of Intelligent Fuzzy Systems*, vol. 37, no. 6, pp. 8535–8550, 2019.
- [48] Z. Deng, K.-S. Choi, Y. Jiang, and S. Wang, "Generalized hidden-mapping ridge regression, knowledge-leveraged inductive transfer learning for neural networks, fuzzy systems and kernel methods," *IEEE Transactions on Cybernetics*, vol. 44, no. 12, pp. 2585–2599, 2014.