

Research Article

A Novel Stress State Assessment Method for College Students Based on EEG

Li Liu ^{1,2}, Yunfeng Ji,¹ Yun Gao,¹ Tao Li,¹ and Wei Xu¹

¹Jiangsu Vocational College of Information Technology, Wuxi, Jiangsu 214153, China

²Jiangsu Key Laboratory of Media Design and Software Technology (Jiangnan University), Wuxi 214122, China

Correspondence should be addressed to Li Liu; liul2@jsit.edu.cn

Received 4 April 2022; Revised 22 May 2022; Accepted 23 May 2022; Published 7 June 2022

Academic Editor: Shengrong Gong

Copyright © 2022 Li Liu 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.

Stress is an unavoidable problem for today's college students. Stress can arouse strong personal emotional and behavioral responses. Compared with other groups of the same age, college students have a special way of life and living environment. They have complex interpersonal relationships and relatively weak social support systems. At the same time, they also face fierce competition in both academic and employment. However, they lack the skills to deal with the crisis and are reluctant to ask others for help, which leads to a simultaneous increase in mental stress. The pressure on college students mainly comes from study, family, social, employment, society, and economy. When students face multiple pressures from family, school, society, etc., some students are prone to some psychological problems due to their own personality or external environment and other reasons. Therefore, regular assessment of students' stress status is an important means to prevent college students' psychological problems. Considering that in real life, the number of students whose pressure is within the tolerable range is the majority, while the number of students who are under too much pressure is a minority. Therefore, the actual dataset to be identified belongs to a kind of imbalanced data. In this study, an improved extreme learning machine (IELM) is used to improve the performance of the recognition model as much as possible. IELM takes the idea of label weighting as the starting point, introduces the AdaBoost algorithm, and combines its weight distribution with the label weighted extreme learning machine (ELM). During the weight update process, the advantage of the imbalanced nature of multi-label datasets is taken. IELM was used to classify EEG data to determine the stress level of college students. The experimental results demonstrate that the algorithm used in this study has excellent classification performance and can accurately assess students' stress levels. The accurate assessment of stress has provided a solid foundation for the development of students' mental health and has significant practical implications.

1. Introduction

College students play an important role in the development and construction of various projects across the country. Their mental health has an impact not only on individual growth and education but also on the country's and society's long-term stability. However, in recent years, college students have experienced psychological crises on occasion, and the subject has grown in prominence, attracting widespread attention from society and university student employees. To address the psychological issues that college students face, we must first comprehend their stress levels. Second, according to the stress condition, the source of psychological stress in college students is precisely evaluated, and the reasons for psychological stress in college students are

thoroughly examined. Finally, corresponding intervention strategies are proposed. This is the main connotation of college students' work. Psychological stress is a person's physiological changes and emotional fluctuations caused by changes in the external environment and the stress response of the body. The transition of college students from high school to university is an important turning point from school to society. There are all kinds of stress involved in adapting to a new environment, learning new courses, facing new challenges, and developing new relationships. The sources of psychological pressure on college students mainly include the following aspects: first, the psychological pressure caused by role change and adaptation disorders; second, the learning pressure caused by the change in learning style and weak learning motivation; third, the interpersonal

pressure caused by communication difficulties and weak communication skills; fourth, the economic pressure and mental pressure caused by the family's economic difficulties; fifth, the employment pressure caused by the severe employment situation and the fierce competition for talents; sixth, the psychological pressure caused by lack of sexual knowledge and immature concept of love; seventh, the psychological pressure caused by personality and emotional problems; and eighth, psychological pressure caused by personality and emotional problems. The psychological stress of college students directly affects their learning effect and quality of life during their school days. For schools, it is about whether they can produce graduates with excellent mental health and professional quality. In the research on the stress of college students, some scholars believe that college students are a high-stress group. Other studies believe that college students only feel less stress, but the sense of stress and the way of coping with stress are indeed the most important factors that cause college students' mental health problems. If psychological pressure cannot be relieved in time, it may lead to high blood pressure and cardiovascular disease and endanger physical health. Depression, pessimism, and hopelessness accompanying psychological pressure will affect students' academic performance and interpersonal communication. This will lead to college students' life satisfaction and happiness. It can even lead to vicious incidents such as wounding and suicide.

Most of the methods of psychological stress assessment use objective scales. Reference [1] compiled a social readjustment scale (SRRS) containing 43 items, which opened the first instance of using questionnaires to measure stress. Subsequently, some researchers have compiled some stress questionnaires with relatively high reliability and validity, such as the Student Stress Scale prepared in [2] and the Graduate Stress Scale prepared in [3]. Commonly used psychological scales are perceived stress scale (PSS) [4], relative stress scale (RSS) [5], psychological stress measure (PSM) [6], and so on. When the method of filling in the scale is used to evaluate the psychological stress state of the subjects, a large-scale test can be administered in a short period of time, and the test results can be obtained quickly. Compared with other methods, it has the advantages of high efficiency and good scientificity. However, many objective scales used in China are compiled based on the psychological characteristics of Westerners. However, the psychological characteristics of social approval, default tendency, and strong conformity in Chinese people's response to the scale have led to a considerable proportion of students' answers with false elements in the scale test. This affects the authenticity and validity of the questionnaire. Therefore, it is difficult to truly evaluate their psychology and behavior through questionnaires. Due to the shortcomings of the scale test method and the complexity of the respondents' motivation to answer the questions, many scales have already added a certain number of fraud identification questions during the compilation process.

To find a more objective stress assessment method, many scholars began to study stress assessment methods based on physiological signals. Studies have shown that there is indeed

a close relationship between psychological stress and brain activity [7]. In particular, when negative emotions appear, the activity of the right hemisphere of the brain is abnormally active [8]. Therefore, scholars have begun to boldly hypothesize that the relative activity of EEG on the right side of the brain predicts changes in psychological stress and a higher risk of mental illness [9]. Based on this conclusion, many EEG-based stress recognition studies have appeared one after another. Reference [10] studied a number of subjects who faced examination pressure and found that the subjects' right forehead EEG activity was more intense under high-intensity examination pressure. Reference [11] studied the stability of prefrontal EEG asymmetry for detecting psychological stress and depression levels, and the results showed that resting EEG α wave asymmetry could be used as a reliable indicator for detecting stress and depression levels. However, some literature studies also pointed out that gender and age differences have a greater impact on the study of psychological stress [12]. Reference [13] proposed that gender differences in prefrontal asymmetry and negative emotion processing may be related to human genes.

To sum up, we found that the existing researches on stress state recognition are based on different types of data such as scales, EEG signals, ECG signals, speech signals, video signals, and facial expressions. The evaluation models used are also different for different data types. Common models are mainly based on machine learning [14–16] and deep learning [17–19]. Because the physiology is more realistic, this study mainly chooses the data based on the physiological signal to identify the stress state. Among the physiological data, EEG and ECG are the most common. Considering that the ECG signal acquisition equipment requires multiple electrodes to collect signals, even a single-lead wearable product is not very convenient. In contrast, EEG signals are easier to acquire, and the signals are sensitive, which can quickly reflect changes in pressure. Therefore, this study chooses EEG as the data type used in the research. In the selection of the recognition model, considering that although the recognition rate of the deep learning algorithm is better, its model training time and high requirements for the hardware performance of the device are high, and the model has a lot of parameters. In this study, the theory is simple and easier to implement the machine learning algorithm. Considering that in real life, the number of students whose pressure is within the tolerable range is the majority, while the number of students who are under too much pressure is a minority. Therefore, the actual dataset to be identified belongs to a kind of imbalanced data. In this study, IELM is used to improve the performance of the recognition model as much as possible. IELM takes the idea of label weighting as the starting point, introduces the AdaBoost algorithm, and combines its weight distribution with the label weighted extreme learning machine. During the weight update process, the advantage of the imbalanced nature of multi-label datasets is taken. IELM was used to classify EEG data in order to determine the stress level of college students. The experimental results demonstrate that the algorithm used in this study has excellent classification performance and can accurately assess students' stress levels.

TABLE 1: Brain wave band details.

Name	Frequency (Hz)	Location	Generated time
δ	0.5 < 4	Forehead in adults, back of brain in children	Occurs mostly in the brains of infants, but also occurs when adults are in deep sleep, coma, or anesthesia
θ	4–7	Brain regions unrelated to hand function	Occurs in young children and adolescents, but also in adults who are tired but conscious
α	8–15	The back of the brain, the resting state is concentrated in the center	Relaxed/contemplative state with eyes closed
β	16–31	The brain is symmetrically distributed on both sides, with a prominent forehead	Positive thinking, focus, vigilance, anxiety
γ	32–45	Somatosensory cortex	Short-term memory, hearing, and touch, multisensory processing

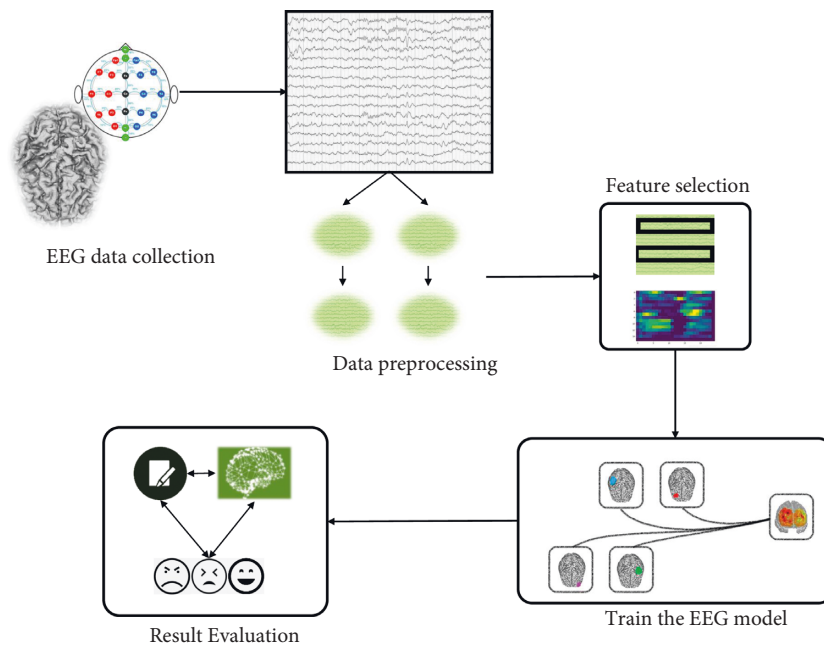


FIGURE 1: EEG-based stress assessment process.

2. EEG-Based Stress Detection

2.1. The Relationship between EEG and Students' Stress State.

Stress, pleasure, focus, etc., are all related to human emotional thinking. The brain controls human thoughts and emotions. When the brain deals with emotion-related issues, the amygdala, orbitofrontal cortex, anterior cingulate cortex, insula, nucleus accumbens, thalamus, and ventral tegmental area of the brain all respond differently to specific emotions. For example, the anterior cingulate cortex responds when we make a decision about something in a happy or sad emotional state. The nucleus accumbens becomes active when we anticipate a reward or something good will happen. When something disgusts us, the insula becomes very active. When these areas are stimulated and activated, neurons in the cerebral cortex are stimulated to generate action potentials. When the cells are stimulated by the impulse, the pyramidal cells will be depolarized, forming a potential difference and generating an electric current. After the electrodes are placed on the scalp, and after amplification, the EEG signal can be collected and an EEG can be drawn. People divide EEG

signals into δ , θ , α , β , and γ frequency bands according to different frequency bands. The descriptions of brain waves in each frequency band are shown in Table 1.

Research has shown that when we blink or think, α waves instantly disappear and reappear, and β waves become very active and have high amplitudes. γ is very sensitive to emotional changes and cognitive learning. When we are tired and sleepy, θ waves begin to appear and become active. To sum up, it can be found that the generation and activity of different brain waves can intuitively reflect the current activity state of students. Therefore, by studying the brain wave signals of these frequency bands, we can analyze the emotional state of college students in the process of study and life and then analyze the stress state of students.

2.2. EEG-Based Stress Assessment. Figure 1 shows the stress assessment process based on EEG signals.

As shown in Figure 1, the first step is to collect EEG data. EEG data collected are all from the freshman to third-year students of our school. Since EEG is a very weak

TABLE 2: Linear and nonlinear characteristics of EEG.

Feature type	Feature name
Linear feature	Full-band center frequency, Hjorth parameter, peak-to-peak, variance, slope, kurtosis
Nonlinear feature	CO complexity, correlation dimension, power spectral entropy, full-band power spectral entropy, Shannon entropy, Kolmogorov entropy

physiological electrical signal, its amplitude can generally only reach the order of microvolts. Therefore, during the acquisition process, various noises generated by different reasons such as the surrounding environment, eye movement signals, and EMG signals are easily mixed into the collected EEG, resulting in poor experimental results. Therefore, to ensure EEG's purity, it is often necessary to perform preprocessing such as denoising the EEG before analyzing the EEG data.

To reduce the data dimension and facilitate subsequent classification processing, feature extraction operations on the preprocessed data are usually required. At present, the research on EEG signal features mainly focuses on linear features, such as extracting the frequency, power, and other features of EEG signals by means of the Fourier transform. However, the structure of the human brain is more complex, and nonlinear features often have better performance. Generally, the linear and nonlinear characteristics are shown in Table 2.

A stress state recognition model is trained based on the EEG training dataset. The algorithm used is the IELM algorithm given in Section 3. The performance of the trained model is validated using the test dataset. The higher the classification accuracy of the sample, the better the model.

3. Imbalanced Extreme Learning Classification Algorithm

3.1. Weighted Extreme Learning Algorithm. For class-balanced tasks, weighted extreme learning is particularly successful. However, it has two drawbacks. First, as the size of the training set grows, so does the time complexity. The second issue is a lack of mistake compensation flexibility. Reference [20] developed a marker-weighted extreme learning machine based on the concept of cost-sensitive learning to address the drawbacks of this technique. By increasing the expected output value of minority class labels, label-weighted extreme learning machines improve the training error tolerance of minority class cases. Furthermore, because it does not use a weight matrix in the optimization method, it has the same time complexity as standard ELM.

$$\begin{aligned}
 L_{ELM} &= \frac{1}{2} \|\beta\|^2 + \gamma \frac{1}{2} W \sum_{i=1}^N \|\tau_i\|^2, \\
 h(x_i)\beta &= t_i^T - \tau_i^T, \\
 i &= 1, 2, \dots, n.
 \end{aligned} \tag{1}$$

Assume that the expected matrix T has m rows and n columns, where m is the number of categories and n is the number of training set samples. The penalty factor is γ . The

output layer weight to be solved is denoted by β . W is an n -dimensional diagonal matrix. Each diagonal element's value corresponds to the penalty factor regulation parameter of the corresponding sample.

$\tau_i = [\tau_{i1}, \tau_{i2}, \dots, \tau_{im}]$ represents the training error vector corresponding to the sample x_i on all output nodes. The expression for T is as follows:

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1N} \\ t_{21} & t_{22} & \cdots & t_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mN} \end{bmatrix}. \tag{2}$$

For the setting of label weights in two-class and multi-class classification problems, [20] provides two weight distribution methods, as follows:

$$\begin{aligned}
 t_{ij} &= \begin{cases} \Delta_{\text{major}}(\text{num})/\Delta(\text{num}_i), & \text{if } x_j \in \text{num}_i, \\ -1, & \text{if } x_j \notin \text{num}_i, \end{cases} \\
 t_{ij} &= \begin{cases} \sqrt[2]{\Delta_{\text{major}}(\text{num})/\Delta(\text{num}_i)}, & \text{if } x_j \in \text{num}_i, \\ -1, & \text{if } x_j \notin \text{num}_i. \end{cases}
 \end{aligned} \tag{3}$$

where $\Delta(\text{num}_i)$ is the number of samples from the i th class in the training set and $\Delta_{\text{major}}(\text{num})$ is the number of samples from the majority class. Two weight distribution methods are provided:

$$\begin{aligned}
 t_{ij} &= \begin{cases} \sim \Delta_{\text{major}}(\text{num})/\Delta(\text{num}_i), & \text{if } x_j \in \text{num}_i, \\ -1, & \text{if } x_j \notin \text{num}_i, \end{cases} \\
 t_{ij} &= \begin{cases} \sqrt[2]{\sim \Delta_{\text{major}}(\text{num})/\Delta(\text{num}_i)}, & \text{if } x_j \in \text{num}_i, \\ -1, & \text{if } x_j \notin \text{num}_i, \end{cases}
 \end{aligned} \tag{4}$$

The weights of the majority class samples remain unchanged in the above weight distribution method, while the weights of the minority class samples are increased. The greater the class imbalance ratio, the greater the weight ratio between the minority and majority classes. The following steps are taken by the weighted extreme learning (Algorithm 1).

3.2. AdaBoost Algorithm. In ensemble learning, according to the different ways of generating base learners, it can be divided into serial ensemble learning algorithms represented by boosting [21] and parallel ensemble learning algorithms represented by bagging [22]. Among the boosting learning paradigms, the AdaBoost algorithm is the most famous, originally proposed by Freund and Schapirel in 1997.

Input: Training set X_1 , test set X_2
Output: Weight β
Step 1: Calculate the number of samples in each category in X_1
Step 2: Create the initial expectation matrix T
Step 3: For x_i in X_1
Step 4: Determine the corresponding t_{ij} in T
Step 5: Calculate t_{ij} according to the above formula
Step 6: Generate hidden layer parameters at random
Step 7: Using the hidden layer parameters and the training set X_1 , compute the hidden layer output matrix H
Step 8: Calculate the output weight β of the weighted extreme learning machine by equation (4)

ALGORITHM 1: LW-ELM algorithm.

Input: Training set X , number of base classifiers M , base classifier H , test sample X'
Output: Class label Y' of test sample X'
Step 1: For $i = 1 : M$
Step 2: Train the base classifier H_i on X_i
Step 3: Calculate the weighted error by equation (5)
Step 4: If the weighted error is greater than 0.5, update all sample weights through equation (7)
Step 5: Calculate the weights of the base classifier H_i by equation (6)
Step 6: Calculate the classification labels of the test samples: $y' = \arg \max_{\sum_{i=1}^n c_i F(H_i(x')) = y}, y \in Y$

ALGORITHM 2: AdaBoost algorithm.

Suppose X is a set containing n training samples for training an AdaBoost classifier. In the AdaBoost algorithm, the importance of each base classifier H_i depends on its error rate ε_i :

$$\varepsilon_i = \frac{1}{n} \left[\sum_{j=1}^n w_j F(H_i(x_j) \neq y_j) \right], \quad (5)$$

where F is the indicator function. If the condition is true, it takes the value 1; otherwise, it is 0. Through the above formula, the importance weight of the base classifier H_i is given as follows:

$$c_i = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_i}{\varepsilon_i} \right). \quad (6)$$

It can be seen that when the error rate c_i is close to 0, $w_i^{(j)}$ corresponds to a large positive value. When the error rate is close to 1, $w_i^{(j)}$ corresponds to a large negative value.

$$\begin{aligned} w_i^{(j+1)} &= \frac{w_i^{(j)}}{S_j} \times e^{-c_j L_j(x_i) y_i} \\ &= \frac{w_i^{(j)}}{S_j} \times \begin{cases} e^{-c_j} & \text{if } H_j(x_i) = y_i, \\ e^{c_j} & \text{if } H_j(x_i) \neq y_i, \end{cases} \end{aligned} \quad (7)$$

where S_j is a normalization factor to ensure that $\sum_i w_i^{(j)} = 1$.

Through the above weight update formula, it is possible to increase the weight of the misclassified samples in the previous round and reduce the weight of those that have been correctly classified. The AdaBoost algorithm's execution steps are as follows (Algorithm 2).

3.3. *IELM*. Reference [23] weights the tokens to increase the expected output of minority class samples. By improving the training error tolerance of the minority class samples, the overall training error of the minority class is similar to that of the majority class in the global training error. In this way, the class imbalance problem is solved. Since the label weight only changes the expected output size and does not change the original optimization formula of the extreme learning machine, the time complexity does not change. However, this algorithm also has certain shortcomings. For example, the algorithm uses human experience to set the weights, which lacks flexibility. The core idea of the AdaBoost algorithm is to change the distribution of samples by modifying the weight of the samples, so that the classifier gradually focuses on those samples that are easy to be misclassified. In this way, the quality of the classification model is maximized. In the label weighting process, larger weights are often assigned to those important samples to avoid misclassification. The above two strategies have the same idea. Based on the inspiration of this idea, this study combines AdaBoost with a labeled weighted extreme learning machine. To efficiently deal with multi-label imbalanced data, the algorithm used improves the two key aspects of AdaBoost's initial weight setting and weight distribution update. The improved model always considers the inherent imbalance characteristics of multi-label and adjusts the size of the weights directionally.

3.3.1. *Adjust the Initial Weight*. For single-label classification problems, the weight distribution reflects the relative importance of the samples. Training samples that are often misclassified tend to receive larger weights than correctly

classified samples. For multi-label classification problems, the weight distribution can reflect the relative importance of the labels. Therefore, the object of weight setting is the marker. The traditional weight setting method is to use the method of evenly distributing the weight. If the data distribution is unbalanced, the weight can be increased by updating the weight in the iterative process and adjusting the minority class adaptively. In fact, the unbalanced degree of the data can be considered. Giving the minority class a higher weight and the majority class a lower weight will inevitably cause the model to converge faster.

Assuming that the multi-label data $S = \{(x_i, Y_i) | i = 1, 2, \dots, n\}$ are a training set containing n samples, $V = \{v_j | 1, 2, \dots, c\}$ represents a label set with c categories, where x_i is the feature vector of the i th sample, and $Y_i \subseteq V$ is its associated label set.

Reference [24] proposed an intra-marker imbalance measure, which has been widely used. For the j th marker, $S_j^+ = \{(x_i, +1) | y_j \in Y_i, 1 \leq i \leq n\}$ represents the positive class sample, $S_j^- = \{(x_i, -1) | y_j \notin Y_i, 1 \leq i \leq n\}$ represents the negative class sample, and the imbalance rate of the j th marker is as follows:

$$R_j = \frac{\max(|S_j^+|, |S_j^-|)}{\min(|S_j^+|, |S_j^-|)}. \quad (8)$$

The initial weight is set as an asymmetric matrix W according to the imbalance ratio, and W contains n rows and c columns, representing n samples and c markers, respectively. The j th token value of the i th sample is as follows:

$$t_{ij} = \begin{cases} \sqrt{(R_j)}, & \text{if } x_i \in \text{num}_j, \\ 1, & \text{if } x_i \in \text{num}_j, \end{cases} \quad (9)$$

$$W_{ij} = \frac{t_{ij}}{p_i}, \quad (10)$$

where p_i is the normalization factor, and its function is to constrain the sum of the weights of the markers to be 1. The square root of R is chosen here because the imbalance ratio in the multi-label is relatively high, and direct use will cause the weight of positive samples to be much larger than the weight of negative samples, making the model fall into the other extreme.

3.3.2. Weight Update. In this study, the purpose of setting asymmetric weight distribution is to make AdaBoost always focus on the imbalance problem within the label. In the iterative process, if the weights are updated in the usual way, the model cannot keep focusing on the multi-label imbalance problem. Therefore, this study updates the weights separately for each category of each token. For the l th label, its j th class error rate is calculated as follows:

$$\varepsilon = \sum_{x_i \in \text{class } j: H(x_i) \neq Y_i^l} S_t(x_i), \quad (11)$$

where H represents the base classifier used by the model and $H(x_i)^l$ represents the output of the base classifier on the l th token of the sample instance x_i .

$$\zeta^{lj} = \frac{1}{2} \ln \left(\frac{\sum_{x_i \in \text{class } j: H(x_i)^l = Y_i^l} S_t(x_i)}{\sum_{x_i \in \text{class } j: H(x_i)^l \neq Y_i^l} S_t(x_i)} \right). \quad (12)$$

When the error rate is close to 0, ζ corresponds to a large positive value. When the error rate is close to 1, ζ corresponds to a large negative value. The calculation formula of the weight W of the $t+1$ th round is as follows:

$$W_{t+1}^{lj}(x_i) = \frac{W_t(x_i) \exp(-\zeta_t^{lj} F(H(x_i), j))}{p_t^{lj}}, \quad (13)$$

where $F(\cdot)$ is an indicator function, and the role of p_t^{lj} is to ensure $\sum W_{t+1}^{lj}(x_i) = 1/2q$ at this time. The weight calculation formula of the entire trainer is as follows:

$$\zeta = \frac{1}{2} \ln \left(\frac{\sum_l \sum_{i: H(x_i)^l = Y_i^l} S_t(x_i)}{\sum_l \sum_{i: H(x_i)^l \neq Y_i^l} S_t(x_i)} \right). \quad (14)$$

3.3.3. IELM Algorithm Execution Steps. The IELM algorithm's execution steps (Algorithm 3) are as follows.

4. Experiment

4.1. Experimental Data Collection. To ensure that the experiment is not affected by factors such as the subject's physical health, other disturbances are minimized as much as possible. Therefore, before this experiment, we learned about the physical condition and basic information of each subject by means of a questionnaire, including the subject's name, age, gender, physical condition, academic status, family status, and the relationship with classmates and teachers. In the questionnaire, subjects are required to check the stress self-evaluation items. The table design of the questionnaire is shown in Table 3.

To better collect EEG without external interference, subjects were first allowed to fall asleep in a quiet environment. When the subject entered deep sleep, the subject's EEG was collected. A total of 35 questionnaire data and EEG data were collected from freshmen to junior college students. These data are simply screened according to the questionnaires filled in by the subjects. For example, when there is a large deviation between the subjects' emotions and the expected emotions of the videos watched, or the subjects are not in the state at all when watching the videos, this set of data will be discarded. Finally, 90 valid samples and 360 EEG fragments were screened out.

4.2. Multifeature Combination Experiment. There are many features of EEG data. This study mainly extracts four features: Hurst index (P1), volatility index, sample entropy, and permutation entropy. The collected EEGs were classified

Input: Training set X , test set X'
Output: Classification evaluation index value
Step 1: Initialize the weight W_1 using equations (9)-(10)
Step 2: For $i = 1 : M$
Step 3: Train base classifier H_i on W_i
Step 4: Calculate the overall error rate for all markers using equation (11)
Step 5: When the weighted error is greater than 0.5, update all sample weights using equation (13), otherwise the loop stops
Step 6: Calculate the weights of the base classifier H_i using equation (14)
Step 7: Calculate the class label of the test sample $y' = \text{sign}(\sum_{i=1}^M c_i H_i(x'))$
Step 8: Calculate the evaluation index value according to the evaluation index calculation formula

ALGORITHM 3: IELM algorithm.

TABLE 3: Stress self-assessment form.

Subject number				
Gender	Male <input type="checkbox"/>		Female <input type="checkbox"/>	
Grade	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	
Study stress	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	
Life pressure	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	
Family stress	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	
Overall pressure	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	

TABLE 4: Classification accuracy based on ELM under four features.

Feature	α	β	γ	δ	θ
Hurst index	0.7122	0.6945	0.7342	0.5898	0.6342
Fluctuation index	0.8098	0.8120	0.7788	0.7556	0.7861
Sample entropy	0.6541	0.5987	0.5062	0.6012	0.5142
Permutation entropy	0.7193	0.6298	0.5865	0.7181	0.5880

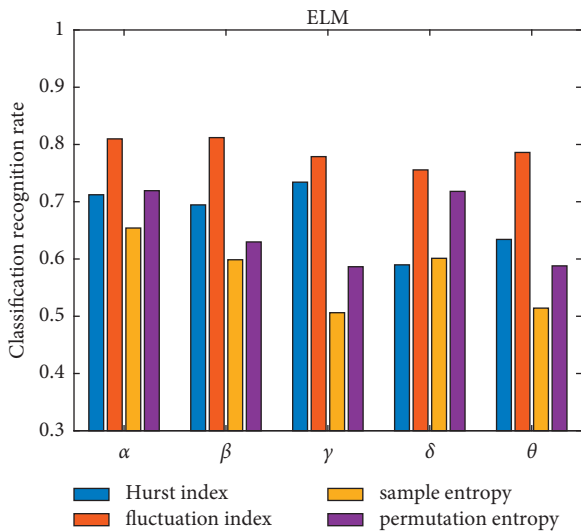


FIGURE 2: Comparison chart of classification accuracy based on ELM under four features.

using radial basis neural network (RBFNN) and IELM. The number of categories is 3, which are high stress, average stress, and low stress. The two algorithms run 10 times on each rhythm to get the average classification accuracy.

TABLE 5: Classification accuracy based on IELM under four features.

Feature	α	β	γ	δ	θ
Hurst index	0.7778	0.7624	0.7997	0.6175	0.7037
Fluctuation index	0.8134	0.8878	0.7959	0.7602	0.8178
Sample entropy	0.7190	0.6530	0.5368	0.6091	0.5392
Permutation entropy	0.7327	0.6490	0.5997	0.7204	0.5914

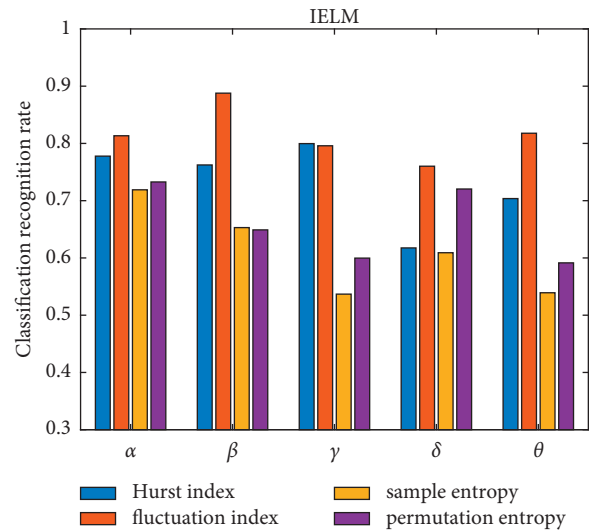


FIGURE 3: Comparison chart of classification accuracy based on IELM under four features.

Table 4 and Figure 2 show the experimental results based on the ELM algorithm. Table 5 and Figure 3 show the experimental results based on the IELM algorithm.

Figures 2 and 3 show that, regardless of classifier, the fluctuation index has the best classification effect, followed by the Hurst index and permutation entropy, and the worst is sample entropy. This demonstrates that different types of features contribute to the classification results at different rates. Figure 4 depicts the classification accuracy of various classifiers from four different perspectives.

As can be seen from Figure 4, under a single feature, the classification accuracy of the IELM algorithm under a single feature is higher than that of the ELM. To resolve the optimal combination of multiple features to improve the classification

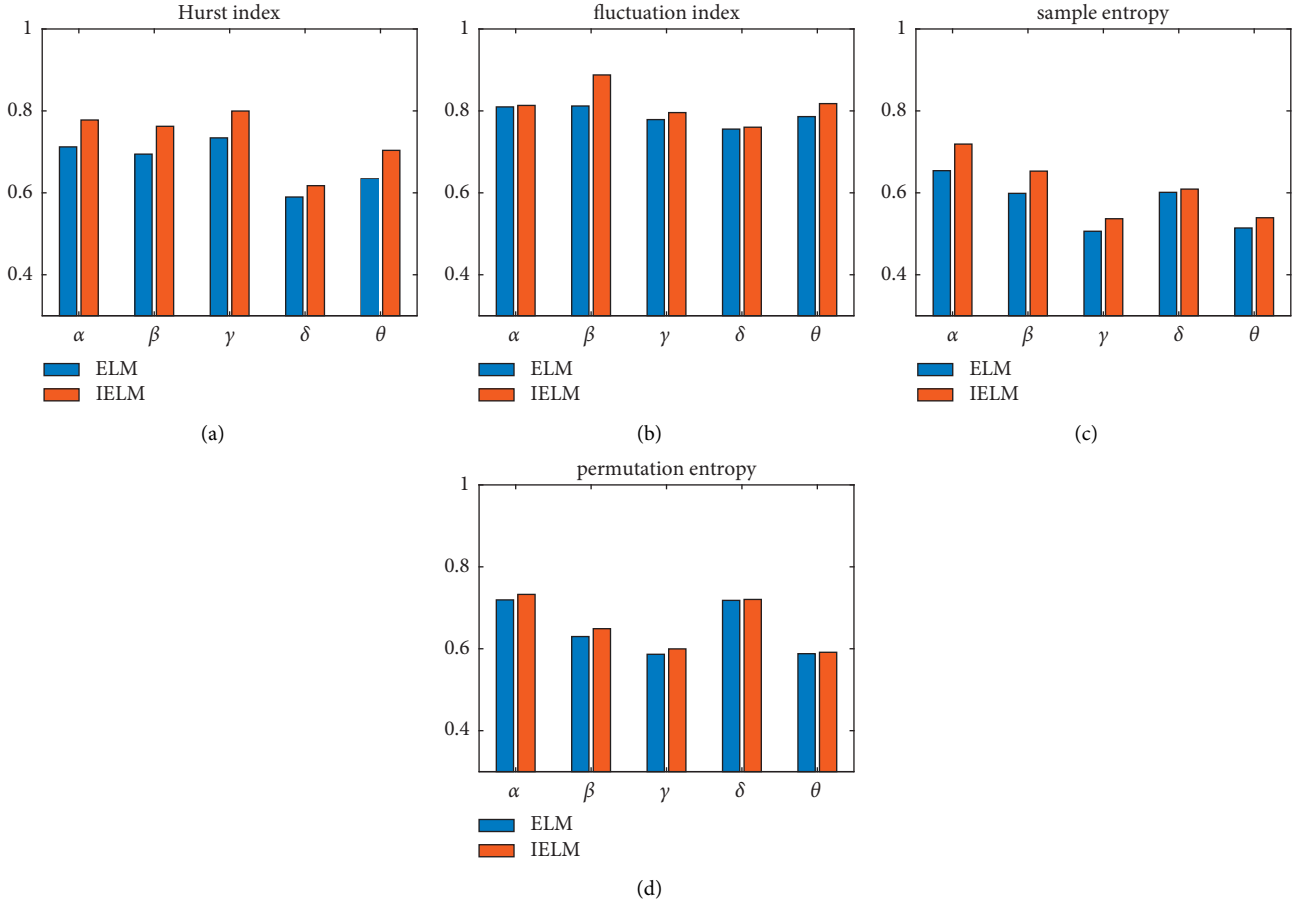


FIGURE 4: Classification accuracy of 4 features. (a) Hurst index. (b) Fluctuation index. (c) Sample entropy. (d) Permutation entropy.

recognition rate of EEG, in this study, a linear weighting formula is given to obtain the weight of each EEG feature. The weight calculation formula is as follows:

$$w_i = \frac{f_i}{f_1 + f_2 + f_3 + f_4}, \quad (15)$$

$$\text{s.t. } 1 \leq i \leq 4, 0 \leq w_i \leq 1,$$

where f_i is the classification recognition rate based on the i th EEG feature and w_i is the weight of the feature f_i . The data in Table 5 are substituted into equation (15), and the calculated feature weights are shown in Table 6.

The weights of the combined features under each frequency band can be obtained from the data in Table 6, and the feature combination formulas under 5 frequency bands are as follows:

$$F_\alpha = 0.2556P1 + 0.2673P2 + 0.2363P3 + 0.2408P4, \quad (16)$$

$$F_\beta = 0.2582P1 + 0.3007P2 + 0.2212P3 + 0.2198P4, \quad (17)$$

$$F_\gamma = 0.2927P1 + 0.2913P2 + 0.1965P3 + 0.2195P4, \quad (18)$$

$$F_\delta = 0.2281P1 + 0.2808P2 + 0.2250P3 + 0.2661P4, \quad (19)$$

TABLE 6: Details of each feature weight.

Feature	α	β	γ	δ	θ
Hurst index	0.2556	0.2582	0.2927	0.2281	0.2653
Fluctuation index	0.2673	0.3007	0.2913	0.2808	0.3084
Sample entropy	0.2363	0.2213	0.1965	0.2250	0.2033
Permutation entropy	0.2408	0.2198	0.2195	0.2661	0.2230

$$F_\theta = 0.2653P1 + 0.3084P2 + 0.2033P3 + 0.2230P4. \quad (20)$$

It can be seen from the experimental data in Table 6 that the weights of the Hurst index and fluctuation index are higher than the remaining two features in the five rhythm species. This shows that these two features show more superior performance in the task of stress assessment. The EEG data of the two feature combinations were classified separately using IELM. Table 7 and Figure 5 give the classification results.

It is clear from the experimental results shown in Table 7 and Figure 5 that for the same classification model, the accuracy of the weighted feature combination is significantly higher than that of the unweighted feature combination. This shows that the introduction of feature weighting strategy can improve the classification performance of EEG. The reason is that different types of features have different

TABLE 7: Classification accuracy under different feature combinations.

Feature combination	α	β	γ	δ	θ
Without weighted feature combination	0.8210	0.8932	0.8090	0.7881	0.8012
With weighted feature combination	0.8340	0.9001	0.8559	0.8293	0.8749

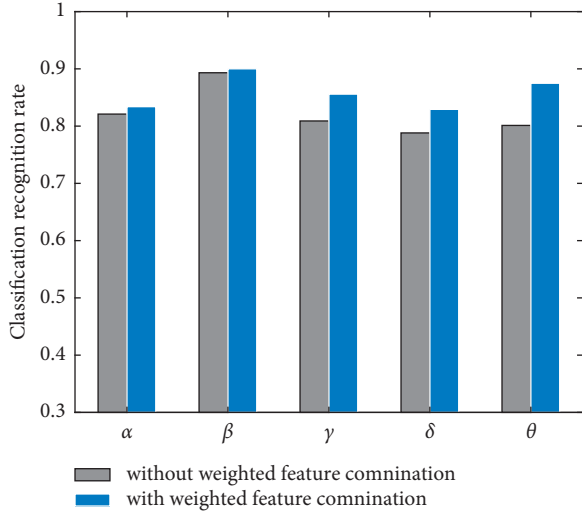


FIGURE 5: Comparison chart of classification accuracy under different feature combinations.

degrees of activity for classification. Based on the weighting formula given in this study, the features with high activity are given a large weight, and the features with low activity are given a small weight, so that the result is optimal.

4.3. Classification Model Experiment. Following the selection of the feature combination method, multiple comparison models are introduced to validate the performance of the classification model used in this study. They are support vector machine (SVM), linear support vector machine (linear SVM), radial basis neural network (RBFNN), random forest (RF), and ELM. Since a single indicator of recognition accuracy cannot fully characterize the performance of the recognition model, this study also introduces the recall indicator to evaluate the model. The experimental results obtained by different classification models are shown in Table 8 and Figure 6.

By observing the experimental results shown in Table 8 and Figure 6, the following experimental conclusions can be obtained:

- (1) Among various classification models, the experimental results obtained by the ELM model are relatively better, close to 0.8. This is one of the reasons why this study chooses ELM as the basic algorithm. When the label weighting and AdaBoost are introduced to optimize the traditional ELM, the experimental results obtained by IELM have been greatly improved. Compared with ELM, its accuracy is improved by 10.25%. This proves the effectiveness of the improved strategy in this study.

TABLE 8: Experimental results of different classification models.

Index/model	SVM	Linear SVM	RBFNN	RF	ELM	IELM
Accuracy	0.6275	0.6598	0.7609	0.7567	0.7912	0.8728
Recall	0.6515	0.6702	0.7122	0.6547	0.7687	0.8913

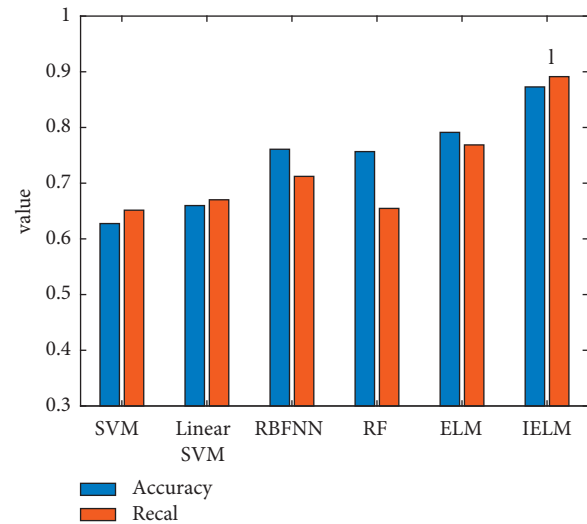


FIGURE 6: Accuracy and recall of different classification models.

- (2) A single indicator is not enough to illustrate the superiority of the model used, so the recall indicator is introduced in the experiment. Compared with SVM, linear SVM, RBFNN, RF, and ELM, the recall rate of the IELM algorithm used is increased by 36.8%, 32.99%, 25.25%, 36.14%, and 15.95%, respectively. The model in this study has been greatly improved in terms of recall rate. This demonstrates the robustness of the algorithm used in this study.

5. Conclusion

Psychological stress has a great impact on human health. The role of early detection and intervention of psychological stress in preventing mental diseases cannot be ignored. However, in terms of stress identification methods, traditional psychological tools such as self-rating scales are highly subjective, and hormone measurement cannot be widely promoted due to its invasiveness and other limitations. The current noninvasive objective evaluation method is still in the research stage. In this study, EEG was selected to assess the psychological stress state of college students. The experiment finally determined the EEG data of 90 students to be used in the experiment. There are 30 students in each grade of freshman, sophomore, and junior. According to the

stress assessment obtained from the questionnaire, 7% of the freshmen were under great stress, 43% were under average stress, and 50% were under little stress. In the sophomore year, 13% were under great stress, 47% were under average stress, and 40% were under less stress. In the third grade, 20% were under great stress, 53% were under average stress, and 27% were under less stress. In general, 14% were under high stress, 47% were under moderate stress, and 39% were under less stress. Among the results obtained by the stress state assessment method based on EEG identification, 12% were under high stress, 46% were under moderate stress, and 42% were under low stress. The stress assessment results obtained using the model used in this study are relatively close to the questionnaire results. This demonstrates that the method described in this study has a certain reference value. However, this study has some limitations. First, whether the self-made questionnaire can collect the most real stress state data needs further research. Second, the number of pressure states used in the experiment is 3, and the division range is relatively general. Future plans are to explore more objective and rigorous data collection methods and optimize the number of stress classifications.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the open project from Jiangsu Key Laboratory of Media Design and Software Technology (Jiangnan University) under grant 21ST0203, Jiangsu Province Higher Vocational Education High-Level Professional Group Construction Project under grant Su Jiaozhi Letter (2021) no. 1, Jiangsu Province Higher Vocational Education Industry-Education Integration Platform Construction Project under grant Su Jiaozhi Letter (2019) no. 26, Natural Science Research Project of Jiangsu Province Colleges and Universities under grant 18KJD510011, Jiangsu Province High-Level Key Professional Construction Project Funding under grant Su Jiaogao (2017) no. 17, and Jiangsu Province Education Science “14th Five-Year Plan” 2021 Annual Project under grant D/2021/03/03.

References

- [1] T. H. Holmes and R. H. Rahe, “The social readjustment rating scale,” *Journal of Psychosomatic Research*, vol. 11, no. 2, pp. 213–218, 1967.
- [2] B. M. Gadzella, “Student-life stress inventory,” *College Students*, vol. 74, p. 12, San Antonio, TX, USA, November 1991.
- [3] I. A. Rochasingh, “Perceived stress among graduate students: development and validation of the graduate stress inventory,” *Educational and Psychological Measurement*, vol. 54, no. 3, pp. 714–727, 1994.
- [4] P. D. Simon, “The 10-item Perceived Stress Scale as a valid measure of stress perception,” *Asia-Pacific Psychiatry*, vol. 13, no. 2, Article ID e12420, 2021.
- [5] I. Ulstein, T. B. Wyller, and K. Engedal, “High score on the Relative Stress Scale, a marker of possible psychiatric disorder in family carers of patients with dementia,” *International Journal of Geriatric Psychiatry*, vol. 22, no. 3, pp. 195–202, 2007.
- [6] L. Lemyre and R. Tessier, “Measuring psychological stress. Concept, model, and measurement instrument in primary care research,” *Canadian family physician Medecin de famille canadien*, vol. 49, no. 9, pp. 1159–1160, 2003.
- [7] Z. Dharmawan, *Analysis of Computer Games Player Stress Level Using EEG data*, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Netherlands, 2007.
- [8] A. N. N. M. Yosi, K. A. Sidek, H. S. Yaacob, M. Othman, and A. Z. Jusoh, “Emotion recognition using electroencephalogram signal,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 15, no. 2, pp. 786–793, 2019.
- [9] R. J. Davidson, D. Pizzagalli, J. B. Nitschke, and K. Putnam, “DEPRESSION: perspectives from affective neuroscience,” *Annual Review of Psychology*, vol. 53, no. 1, pp. 545–574, 2002.
- [10] S. Richard, N. Y. W. Lewis, and T. H. Wang, “The effect of a naturalistic stressor on frontal EEG asymmetry, stress, and health,” *Biological Psychology*, vol. 75, no. 3, pp. 239–247, 2007.
- [11] J. J. B. Allen, H. L. Urry, S. K. Hitt, and J. A. Coan, “The stability of resting frontal electroencephalographic asymmetry in depression,” *Psychophysiology*, vol. 41, no. 2, pp. 269–280, 2004.
- [12] S. D. Ja, D. Posthuma, D. I. Boomsma, and E. J. C. de Geus, “The relation between frontal EEG asymmetry and the risk for anxiety and depression,” *Biological Psychology*, vol. 74, no. 1, pp. 26–33, 2007.
- [13] J. A. Coan, *The Heritability of Trait Frontal EEG Asymmetry and Negative Emotionality: Sex Differences and Genetic Non-additivity*, The University Of Arizona, Tucson, AZ, USA, 2003.
- [14] L. Liu, L. Kuang, and Y. Ji, “Multimodal MRI brain tumor image segmentation using sparse subspace clustering algorithm,” *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 8620403, 13 pages, 2020.
- [15] L. Liu, Y. Ji, Y. Gao, and L. Kuang, “Classification of epileptic electroencephalograms signal based on integrated radius-basis-function neural-network,” *Journal of Medical Imaging and Health Informatics*, vol. 8, no. 7, pp. 1462–1467, 2018.
- [16] A. Özbeyaz, “EEG-Based classification of branded and unbranded stimuli associating with smartphone products: comparison of several machine learning algorithms,” *Neural Computing & Applications*, vol. 33, no. 9, pp. 4579–4593, 2021.
- [17] G. Altan, A. Yayık, and Y. Kutlu, “Deep learning with ConvNet predicts imagery tasks through EEG,” *Neural Processing Letters*, vol. 53, no. 4, pp. 2917–2932, 2021.
- [18] J. R. McIntosh, J. Yao, L. Hong, J. Faller, and P. Sajda, “Ballistocardiogram artifact reduction in simultaneous EEG-fMRI using deep learning,” *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 1, pp. 78–89, 2021.
- [19] L. C. Sarmiento, S. Villamizar, O. López, A. C. Collazos, J. Sarmiento, and J. B. Rodríguez, “Recognition of EEG signals from imagined vowels using deep learning methods,” *Sensors*, vol. 21, no. 19, p. 6503, 2021.
- [20] 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.
- [21] I. F. Kilincer, F. Ertam, and A. Sengur, “A comprehensive intrusion detection framework using boosting algorithms,” *Computers & Electrical Engineering*, vol. 100, Article ID 107869, 2022.
- [22] A. Zahirah, A. B. Selamat, H. Fujita, and O. Krejcar, “The best ensemble learner of bagged tree Algorithm for student performance prediction,” *Knowledge Innovation Through Intelligent Software Methodologies, Tools and Techniques*, vol. 327, pp. 55–64, 2020.
- [23] W. Samek, F. C. Meinecke, and K. R. Muller, “Transferring subspaces between subjects inBrain-computer interfacing,” *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 8, pp. 2289–2298, 2013.
- [24] M. L. Zhang, Y. K. Li, and X. Y. Liu, “Towards class-imbalance aware multi-label learning,” in *Proceedings of the International Conference on Artificial Intelligence*, pp. 4041–4047, Nanjing, China, August 2015.