*Research Article*

# An Automatic Epilepsy Detection Method Based on Improved Inductive Transfer Learning

**Yufeng Yao** [1,2] **and Zhiming Cui** [3]

$^1$*The Institute of Intelligent Information Processing and Application, Soochow University, Suzhou 215006, China*
$^2$*Department of Computer Science and Engineering, Changshu Institute of Technology, Changshu 215500, China*
$^3$*Suzhou University of Science and Technology, Suzhou 215009, China*

Correspondence should be addressed to Yufeng Yao; yyf@cslg.edu.cn and Zhiming Cui; zmcui@usts.edu.cn

Epilepsy is a chronic disease caused by sudden abnormal discharge of brain neurons, causing transient brain dysfunction. The seizures of epilepsy have the characteristics of being sudden and repetitive, which has seriously endangered patients' health, cognition, etc. In the current condition, EEG plays a vital role in the diagnosis, judgment, and qualitative location of epilepsy among the clinical diagnosis of various epileptic seizures and is an indispensable means of detection. The study of the EEG signals of patients with epilepsy can provide a strong basis and useful information for in-depth understanding of its pathogenesis. Although, intelligent classification technologies based on machine learning have been widely used to the classification of epilepsy EEG signals and show the effectiveness. In fact, it is difficult to ensure that there is always enough EEG data available for training the model in real life, which will affect the performance of the algorithms. In view of this, to reduce the impact of insufficient data on the detection performance of the algorithms, a novel discriminate least squares regression-(DLSR-) based inductive transfer learning method was introduced which is on the basis of DLSR and the inductive transfer learning. And, it is applied to promote the adaptability and accuracy of the epilepsy EEG signal recognition. The proposed method inherits the advantages of DLSR; it can be more suitable for classification scenarios by expanding the interval between different classes. Meanwhile, it can simultaneously use the data of the target domain and the knowledge of the source domain, which is helpful for getting better performance. The results show that the improved method has more advantages in EEG signal recognition comparing to several other representative methods.

## 1. Introduction

At present, epilepsy has become a common disease in neurology. Its pathogenesis has not yet been fully elucidated, and it is usually defined as a chronic neurological disease caused by sudden abnormal discharge of brain neurons. The epileptic seizures are sudden and repetitive. Its onset is accompanied by clinical manifestations such as loss of consciousness, fainting, and twitching of extremities. It also has cognitive and mental disorders that seriously endanger patients' health, cognition, etc. [1, 2]. According to statistics, more than one percent of the world's population suffer from the disease [3], and there are approximately 9 million people with epi-

lepsy in China. Therefore, the depth research and prevention of epilepsy play an indispensable role to alleviate the suffering of patients, improve the quality of life, and promote healthy development. As an important method for studying epilepsy, EEG uses electrodes to record the electrical activity of nerve cells in the brain, which contains a large amount of physiological and pathological information, and is of great importance in the clinical examination, location, and therapy of epilepsy. Therefore, for people with a tendency to epilepsy, automatic detection of epilepsy can analyze and screen the EEG signals of people at high risk for epilepsy, so as to realize early detection, perform timely intervention, and reduce the impact of epilepsy on people and the incidence of epilepsy.

In view of this, it is of great value to study the epilepsy automatic detection algorithm based on EEG signals and develop an efficient and accurate epilepsy automatic detection system.

In fact, the study of automatic epileptic detection based on EEG signals has attracted extensive attention from scholars and experts at home and abroad since the 1970s. To predict the onset or preonset of epilepsy in the process of seizure detection, machine learning and pattern classification algorithms are generally applied to classify the EEG signals after extracting the characteristics of the time domain, frequency domain, time frequency domain, or nonlinear domain of the EEG. With the development of computer technology and digital signal processing technology, more and more methods are widely used in the study of seizure detection methods and have achieved certain research results, such as the Bayesian classifier [4], artificial neural network [5–9], support vector machine (SVM) [10–13], and fuzzy reasoning [14, 15]. For example, Obeyli extracted the Lyapunov exponential features of EEG signals and used probabilistic neural networks to classify EEG signals, so as to achieve high classification results [9]. Chan et al. extracted the time-frequency features of five subbands in the wavelet transform domain of epilepsy EEG signals and then used support vector machines and cluster regression models to recognize the onset of seizures. Aarabi et al. [14] extracted the features such as sample entropy, dominant frequency, average amplitude, and amplitude variation coefficient of intracranial EEG da4ta of patients with epilepsy and used the established fuzzy inference rules to fuse the EEG feature information for seizure detection. Although many of the above intelligent classification methods have shown the effectiveness of epilepsy EEG signal classification, they still face a challenge, that is, it is very hard to get enough EEG data for epilepsy to train the model in real life. Therefore, it has important practical value to explore how to use the knowledge acquired from related fields to enhance the classification performance of EEG data in the current scenario [16].

To solve the above challenges, a novel inductive transfer learning method based on discriminant least squares regression (TDLSR) was proposed. Meanwhile, it was applied to specific medical application scenarios, namely, epilepsy EEG signal classification, so as to relieve the effect of severe data shortage on the performance of the algorithms. Transfer learning is an effective way to transfer knowledge from related fields and is helpful of obtaining more information in the absence of sufficient data or information. It focuses on how to use the useful information from similar but different source domains to improve the classification result of the classifier in the target domain. When studying epilepsy EEG signal classification, the inductive transfer learning method naturally becomes the first choice because of insufficient labeled epilepsy EEG samples in the target domain sometimes. What is more, since the discriminant LSR is still based on least squares regression (LSR) [17], which can explain the importance of each feature in the prediction model based on the original data space, we introduced the inductive transfer learning method based on DLSR to use for epilepsy EEG sig-

nal classification. In summary, the innovations of this work are summarized as follows:

*Point 1.* The improved method is on the basis of the inductive transfer learning, but it has some difference from the traditional inductive transfer learning. The latter directly transfers the samples or features used in the source domain to the target domain for transfer learning, while the former uses a knowledge lever mechanism that transfers some knowledge from the source domain to the target domain. Then, the security of the data in the source domain can be well protected, and the data in the target domain and the knowledge in the source domain can be used simultaneously, so that the classification effect is better.

*Point 2.* The improved method expands DLSR that can be well applied to classification scenes into a novel method with certain transfer learning ability, so that it can be used in more complex scenes.

*Point 3.* The improved method inherits the characteristics of DLSR, that is, it can be better applied to classification scenarios by expanding the interval between different categories. And, it can transfer knowledge from the source domain, thus ensuring the rationality of its training model.

Finally, to better illustrate the basic idea of this study, the structure of the paper is as follows:

*Part 1.* Introduction to the research background, status, and significance of the thesis.

*Part 2.* The related work, including the related technology of epilepsy EEG signal detection and the epilepsy EEG signal classification based on transfer learning, is summarized in advance so that the following sections become more readable.

*Part 3.* The notations of the inductive transfer learning algorithm based on DLSR was introduced in detail.

*Part 4.* The reliability and validity of TDLSR algorithm in detection of epilepsy EEG signals based on a series of experimental were verified.

## 2. Related Work

Automatic epilepsy detection is based on signal processing technology and pattern recognition. It analyzes EEG data to identify the location and duration of seizures. Usually, the EEG signals collected during seizures was called seizure EEG, and the EEG signals collected in the nonseizure are called nonseizure EEG. The problem of automatic detection of epilepsy EEG is to effectively judge the above two types of EEG signals and identify seizures. The related detection technologies are introduced as follows.

*2.1. The Related Technology of Epilepsy EEG Signal Detection.* Because the EEG signal of epilepsy is easily interfered by many factors, it is very random, and it is a nonstationary signal, and its rule is generally difficult to grasp. Therefore, researchers often use quantitative analysis to extract characteristic information of epilepsy EEG signals. The existing methods of automatic epilepsy detection include the following:

(1) Time-domain analysis: time-domain analysis is one of the earliest methods used in signal analysis. It

analyzes the time-domain waveforms of EEG signals to study the difference between EEG waveform during seizures and EEG waveform during nonseizures, and directly extracts the waveform characteristics of the signals to distinguish the two types of EEG signals. The time-domain analysis method has the characteristics of intuitiveness and clear physical meaning, and it can reflect the important information in transient waveform such as spike wave and harp wave. The representative methods of time-domain analysis have a template matching method, time-domain waveform, or energy characteristics, etc. In 2001, Litt et al. [18] performed feature extraction on EEG signals based on time-domain analysis methods. In addition, researchers performed time-domain analysis of EEG signals to extract waveforms or energy features different from background activities and use amplitude, rhythm, period, and other parameters as classification criteria to identify epilepsy EEG signals. Gotman et al. performed a "half-wave" decomposition of EEG signals and then extracted EEG features, including average amplitude, duration, and coefficient of variation relative to the background, and set thresholds based on expert experience. The characteristic parameters are compared with the threshold to judge whether it is an epilepsy signal [19–21]. Time-domain analysis only performs statistical analysis from the time domain, and it is easy to miss other important changes in abnormal signals, such as slow waves

(2) Frequency analysis: unlike time-domain analysis, which mainly analyzes the waveform characteristics of epilepsy EEG, frequency domain analysis analyzes the frequency characteristics of EEG. It recognizes different rhythms according to the frequency of brain waves. Each brain wave of different rhythms corresponds to epilepsy EEG signals in different time periods or different parts of the brain [22]. Frequency domain analysis is based on the Fourier transform and is mainly used for power spectrum analysis of EEG signals. It performs the Fourier transform on the EEG signal to obtain its frequency components and spectrum distribution and extracts the corresponding EEG features in the frequency domain for epilepsy detection and recognition. Representative methods include power spectrum estimation, autoregressive (AR) model spectrum estimation, and higher order spectrum [23]. Among them, the power spectrum estimation transforms the EEG signal whose amplitude changes with time into the EEG spectrum chart with power varying with frequency and analyzes the distribution and change of each frequency band of the EEG signal intuitively and quantitatively [24, 25]. Although frequency domain analysis can provide a lot of effective information, allowing researchers to detect epilepsy based on the frequency domain characteristics of EEG, the overall spectrum of the signal obtained by the Fourier transform nei-

ther can reflect the local characteristics of the signal nor can reflect the signal frequency component changes with time. Therefore, the detection results obtained by frequency domain analysis are not very satisfactory and are greatly restricted in practical applications

(3) Time-frequency analysis: the epilepsy EEG is a typical nonstationary signal, which contains not only the waveform parameter characteristics in the time domain but also the energy distribution characteristics in the frequency domain. However, neither the above two methods can fully extract the transient characteristics and information of the EEG signals and can get the ideal results. With the development of digital signal theory and methods, the time-frequency analysis method combining time domain and frequency domain is widely used in the analysis of nonstationary EEG signals. It can obtain time and frequency domain information at the same time and capture transient information of EEG. In recent years, more and more studies have adopted time-frequency analysis methods to analyze EEG signals, among which various wavelet change methods are represented. The wavelet transform uses the translation and expansion of the window function to implement a wide time window for the low-frequency components of the signal and a narrow time window for the high-frequency components to complete the multiscale analysis of the signal. This analysis method conforms to the laws of nature and has a good ability to characterize the local characteristics of the signal [26]. It can capture the transient characteristics of the EEG signal and accurately locate it in the time and frequency domains. In addition to wavelet transform, commonly used time-frequency analysis methods also include empirical mode decomposition [27–30], the Wigner-Ville distribution [31, 32], and the Stockwell transform [33, 34]. However, most of these time-frequency analysis methods can only be used for multiresolution analysis of the original signal and then need to be combined with other algorithms to achieve the feature extraction and selection of EEG. Figure 1 shows the comparison of time-domain analysis, frequency domain analysis, and wavelet transform analysis.

Observe the above figures, it is easy to draw a conclusion that the time-frequency analysis can provide more useful information compared to time-domain analysis and frequency domain analysis. In Figure 1(c), the wavelet transform improves the time resolution at the high frequency of the signal by changing the time window and improves the frequency resolution at the low frequency, which has a better classification effect

(4) Nonlinear dynamic analysis: nowadays, with the progress of nonlinear dynamics theory, researchers are devoted to studying the nonlinear of EEG signals
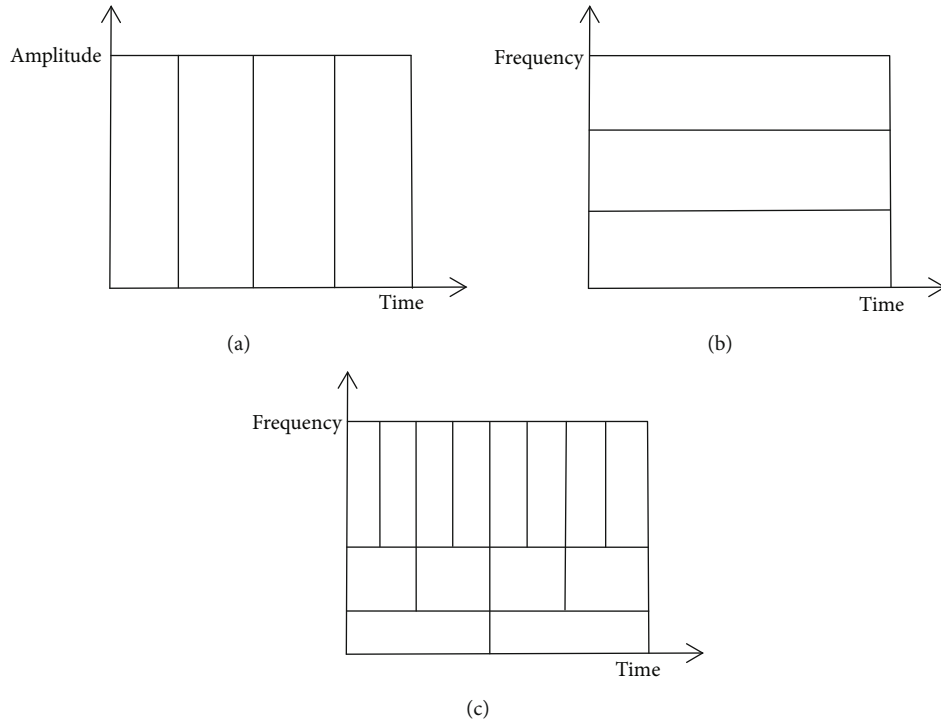
FIGURE 1: Comparison of time-domain analysis, frequency domain analysis, and time-frequency analysis. (a) Time-domain analysis. (b) Frequency analysis (Fourier transform). (c) Wavelet transform.

to solve the problem of automatic detection of epilepsy. Using nonlinear dynamics theory methods for EEG signal analysis, various nonlinear features of EEG signals can be extracted to distinguish epilepsy EEG signals from normal EEG signals. This provides some new research ideas for automatic epilepsy detection technology. Kannathal et al. used different entropies to measure the chaotic characteristics of EEG signals and used them as EEG features to distinguish EEG signals in different periods [35], including the Shannon entropy, Renyi entropy, Kolmogorov-Sinai entropy, and approximate entropy. The results show that the complexity of the EEG signal in patients with epilepsy during the intermittent period is higher than that during the seizure period, that is, the complexity of the EEG signal during the seizure is reduced, and the bet value is less than the normal EEG signal. Although nonlinear analysis can reflect the dynamic mechanism of seizures well, most of the nonlinear features are computationally intensive and generally time-consuming, which is not suitable for real-time epilepsy automatic detection systems.

*2.2. The Epilepsy EEG Signal Classification Based on Transfer Learning.* Traditional classification methods use a large amount of data with label information to train a decision function and then use this function to classify and identify test samples with unknown label information. However, these classification methods all have a presupposition: training data and test data need to obey the same distribution characteristics, as shown in Figure 2. For the differences in
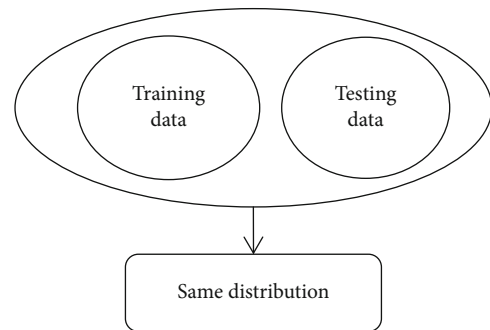


FIGURE 2: Scene suitable for the traditional method.

the distribution of training samples and test samples as described above, the performance of the traditional methods significantly decrease, as shown in Figure 3. In response to this challenge, transfer learning is a promising research direction. Transfer learning focuses on knowledge transfer problems that are similar to different domains or have different data distributions. It enhances the performance of the classifier used for target area recognition by learning useful knowledge from the source domain. According to whether the target domain used contains samples with labeled information, transfer learning techniques are divided into three categories: inductive transfer learning method, direct transductive transfer learning method, and unsupervised transfer learning method [36].. In the paper, we will focus on an inductive transfer learning method with good performance, that is, inductive transfer learning method
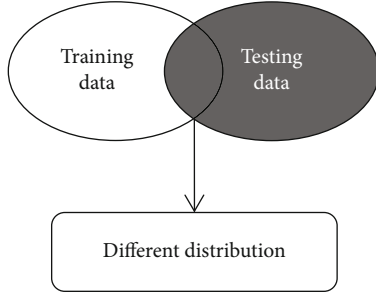
FIGURE 3: New challenge for the traditional methods.

based on discriminant least squares regression (TDLSR). And, its application and actual effect in EEG signal detection of epilepsy will be studied. The framework structure of epilepsy EEG signal detection based on transfer learning theory is given, as shown in Figure 4.

In short, transfer learning is to transfer the knowledge (useful knowledge) from the source domain with a large amount of labeled data for learning in the target domain with no or little labeled data, thereby improving the training quality of the target domain. This can reduce the workload of collecting labeled data in the target domain.

## 3. The Inductive Transfer Learning Algorithm Based on DLSR

To better describe the algorithm proposed in this paper, Table 1 gives a detailed description of the symbols in the algorithm.

Since DLSR is a nontransforming algorithm, the symbols in Table 1 refer to the parameter variables of the original training sample.

*3.1. The Least Squares Regression.* As a widely used method based on statistical theory, LSR has become a typical method. LSR uses the Frobenius norm to constrain the matrix of representation coefficients. In the paper, to expand the classification ability of the LSR algorithm, we preconstructed the binary label matrix $Y$ corresponding to the training samples $X$, so that it can better cope with more complex classification scenarios. The $j$th column of $Y$ indicates that only the data belonging to the $j$th class corresponds to an element equal to 1, and all other elements are 0. Then, the objective function of LSR can be redefined as follows:

$$\arg \min_Z \|XZ - Y\|_F^2 + \lambda \|Z\|_F^2. \tag{1}$$

Since the LSR has an analytical solution, which can be easily obtained, formula (1) can be rewritten as:

$$J(Z) = \arg \min_Z \|XZ - Y\|_F^2 + \lambda \|Z\|_F^2. \tag{2}$$
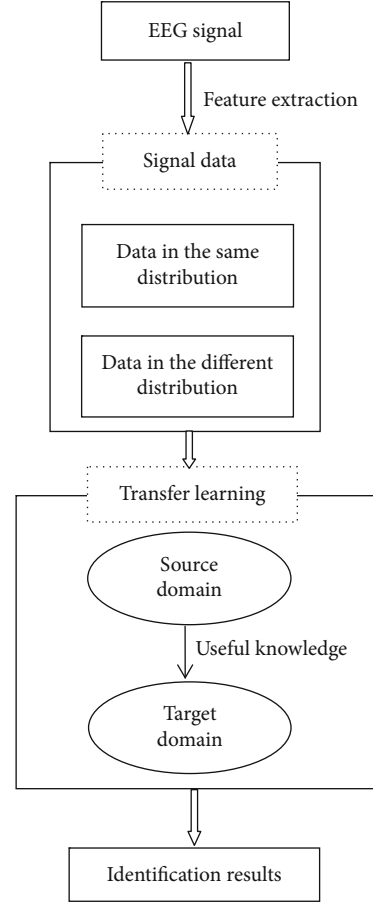


FIGURE 4: Framework of adaptive recognition for epileptic EEG signals based on transfer learning.

TABLE 1: Symbol description.

| Symbol | Description |
|---|---|
| $X$ | Training matrix of the target domain, $X = [x_1, x_2, \cdots, x_N]$ and $x_i \in R^d$ |
| $N$ | Total number of training samples of the target domain |
| $d$ | Dimension of sample |
| $C$ | The number of classes |
| $Y$ | The binary label matrix corresponding to the training samples in the target domain. The $j$th column of $Y$ indicates that only the data belonging to the $j$th class corresponds to an element equal to 1, and all other elements are 0 |
| $B$ | Indicator matrix constructed based on $Y$ |
| $W$ | Label offset matrix of the target domain |
| $Z$ | Mapping matrix of target domain |
| $Z_S$ | Mapping matrix of source domain |
| $e$ | Vector representing all 1, $e \in R^N$ |
| $\Theta$ | Hadamard product operator for matrix |
| $p$ | The mapping offset vector of target domain |
| $\lambda, \eta$ | The regularization parameter |

Then, the solution process of the analytical solution of LSR is as follows: let

$$\frac{\partial J(Z)}{\partial Z} = 0,$$
$$\Longrightarrow 2X^{\mathrm{T}}XZ - 2X^{\mathrm{T}}Y + 2\lambda Z = 0, \tag{3}$$
$$\Longrightarrow (X^{\mathrm{T}}X + \lambda I_d)Z = X^{\mathrm{T}}Y.$$

Finally, the analytical solution of LSR can be obtained as

$$Z = (X^{\mathrm{T}}X + \lambda I_d)^{-1}X^{\mathrm{T}}Y. \tag{4}$$

*3.2. The Discriminate Least Squares Regression.* As we know, LSR can be directly used for classification tasks. However, since the interval between any two different classes in the constructed binary class label matrix is $\sqrt{2}$, the DLSR proposed in literature [37] introduces the relaxation technique into the LSR so as to expand the interval between the two data from different classes. To improve the compactness of the classification task, DLSR will comprehensively consider the class factors and build an indicator matrix $B$ on the basis of the binary label matrix $Y$ of the sample. Each element of matrix $B$ is defined as follows:

$$B_{ij} = \begin{cases} +1, \text{ if } y_i = 1, \\ -1, \text{ otherwise.} \end{cases} \tag{5}$$

In essence, each element of matrix $B$ represents the offset direction of the corresponding class label. Then, $\varepsilon$ relaxation on each element of $Y$ is performed, and the amount of $\varepsilon$ relaxation through the matrix $W$ is recorded. Then, the objective function of DLSR can be expressed as

$$\min_{s.t.W\geq 0} \left\| XZ + ep^{\mathrm{T}} - (Y + B\Theta W) \right\|_{\mathrm{F}}^2 + \lambda \|Z\|_{\mathrm{F}}^2. \tag{6}$$

The objective function of DLSR is a convex optimization problem. However, it cannot directly optimize the solution. Literature [37] adopts an alternative optimization strategy and ensures that a closed solution is obtained at each step. The specific derivation process is as follows:

(1) Fix $W$, and update $Z$ and $p$. Let $L = Y + B\Theta W$; formula (6) can be rewritten as

$$J(Z,p) = \arg\min_{Z,p} \left\| XZ + ep^{\mathrm{T}} - L \right\|_{\mathrm{F}}^2 + \lambda \|Z\|_{\mathrm{F}}^2. \tag{7}$$

According to the optimization theory, we take a partial derivative of $p$, namely

$$\frac{\partial J(Z,p)}{\partial p} = 0,$$
$$\Longrightarrow Z^{\mathrm{T}}X^{\mathrm{T}}e + pe^{\mathrm{T}}e - L^{\mathrm{T}}e = 0, \tag{8}$$
$$\Longrightarrow p = \frac{(L^{\mathrm{T}}e - Z^{\mathrm{T}}X^{\mathrm{T}}e)}{N}.$$

Furthermore, we find the partial derivative of $Z$, and we can get

$$\frac{\partial J(Z)}{\partial Z} = 0,$$
$$\Longrightarrow 2X^{\mathrm{T}}XZ - 2X^{\mathrm{T}}(ep^{\mathrm{T}} - L) + 2\lambda Z = 0,$$
$$\Longrightarrow X^{\mathrm{T}}XZ - X^{\mathrm{T}}\left(e\frac{(e^{\mathrm{T}}L - e^{\mathrm{T}}XZ)}{N} - L\right) + \lambda Z = 0,$$
$$\Longrightarrow X^{\mathrm{T}}\left(I_N - \frac{ee^{\mathrm{T}}}{N}\right)XZ - X^{\mathrm{T}}\left(I_N - \frac{ee^{\mathrm{T}}}{N}\right)L + \lambda Z = 0,$$
$$\Longrightarrow Z = (X^{\mathrm{T}}HX + \lambda I_d)^{-1}X^{\mathrm{T}}HL \tag{9}$$

(2) Fix $Z$ and $p$, and update $W$. Let $G = XZ + ep^{\mathrm{T}} - Y$; then, $W$ can be solved as follows:

$$\begin{aligned} \arg\min_{W} &\|G - B\Theta W\|_{\mathrm{F}}^2. \\ s.t.&W\geq 0 \end{aligned} \tag{10}$$

According to literature [37], the Frobenius norm square of a matrix can be solved element by element, so Eq. (10) can be equivalent to solve $N \times C$ subproblems. For element $W_{ij}$ in row $i$ and column $j$, there can be

$$\min_{W_{ij}} (G_{ij} - B_{ij}W_{ij})^2, \tag{11}$$
$$s.t.W_{ij}\geq 0$$

where $G_{ij}$ and $B_{ij}$ represent the $j$th element of the $i$th row of the matrix $G$ and matrix $B$, respectively and satisfy $B_{ij}^{\ 2} = 1$. Then, we can get

$$(G_{ij} - B_{ij}W_{ij})^2 = (B_{ij}G_{ij} - W_{ij})^2. \tag{12}$$

And, because each element of $W$ satisfies $W_{ij} \geq 0$, formula (12) can be written as

$$W_{ij} = \max(B_{ij}G_{ij}, 0). \tag{13}$$

---

The description of DLSR
**Input:**
The training samples $X = [x_1, x_2, \cdots x_N] \in R^{N \times d}$, and their corresponding class labels $y_i \in \{1, 2, \cdots C\} (i = 1, 2, \cdots, N)$, where $x_i \in R^d (i = 1, 2, \cdots, N)$. The maximum number of iterations is $T$.
**Output:**
The mapping matrix $Z$ and the translation vector $p$.
**Training:**
Step 1: Construct the label matrix $Y$ and the indicator matrix $B$, respectively.
Step 2: Get the value of $\lambda$ by grid searching in the set of $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10, 10^2\}$.
Step 3: Initialize $Z' = 0$, $p = 0$, and $W = 0$. Set $t = 1$.
Step 4: Update Q by $Q = (X^T H X + \lambda I_d)^{-1} X^T H$.
Repeat
Step 5: Update the label matrix $L$ by $L = Y + B \Theta W$.
Step 6: Update $p$ and $Z$ by Eqs. (8) and (10).
Step 7: Update the regression error matrix $G$ by $G = XZ + ep^T - Y$.
Step 8: Update the label shift matrix $W$ by Eq. (14).
Step 9: Let $Z' = Z$ and $t = t + 1$
Until $\|Z' - Z\|_F^2 \leq 10^{-4}$ or $t > T$.
Step 10: Output $Z$ and $p$.

ALGORITHM 1: The description of the DLSR algorithm.

---

The description of TDLSR
**Input:**
The training samples in target domain $X = [x_1, x_2, \cdots x_N] \in R^{N \times d}$, and their corresponding class labels $y_i \in \{1, 2, \cdots C\} (i = 1, 2, \cdots, N)$, where $x_i \in R^d (i = 1, 2, \cdots, N)$. The maximum number of iterations is $T$. $A_S$ is learned in advanced by using DLSR from source domain.
**Output:**
The mapping matrix $Z$ and the translation vector $p$.
**Training:**
Step 1: Construct the label matrix $Y$ and the indicator matrix $B$, respectively.
Step 2: Get the value of $\lambda$ and $\eta$ by grid searching in the set of $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10, 10^2\}$, respectively.
Step 3: Initialize $Z' = 0$, $p = 0$, and $W = 0$. Set $t = 1$.
Step 4: Compute Q by $Q = (X^T H X + \lambda I_d)^{-1} X^T H$ and V by $V = \eta (X^T H X + (\lambda + \eta) I_d)^{-1} Z_S$, respectively.
Repeat
Step 5: Update the label matrix $L$ by $L = Y + B \Theta W$.
Step 6: Update $p$ and $Z$ by Eqs. (17) (18).
Step 7: Update the regression error matrix $G$ by $G = XZ + ep^T - Y$.
Step 8: Update the label shift matrix $W$ by Eq. (20).
Step 9: Let $Z' = Z$ and $t = t + 1$
Until $\|Z' - Z\|_F^2 \leq 10^{-4}$ or $t > T$.
Step 10: Output $Z$ and $p$.

ALGORITHM 2: The description of the TDLSR algorithm.

---

Therefore, the final solution formula of $W$ is:

$$W = \max (B \Theta G, 0). \qquad (14)$$

According to the above derivation, Algorithm 1 gives a detailed description of the DLSR algorithm.

### 3.3. The Inductive Transfer Learning Algorithm Based on DLSR.
Most inductive transfer learning algorithms are implemented by directly learning from the data in the source domain through some classes. However, in the paper, we used a knowledge-based inductive transfer learning framework instead of raw data to study inductive transfer learning methods based on source domain knowledge. Inspired by this, an inductive transfer learning algorithm based on DLSR was introduced. Its objective function is

$$\min_{Z,p} \left\| XZ + ep^T - (Y + B \Theta W) \right\|_F^2 + \lambda \|Z\|_F^2 + \eta \|Z - Z_s\|^2.$$
$$s.t. W \geq 0$$
$$(15)$$

It can be found from the formula (15) that the first two items directly inherit the DLSR for learning in the target domain. The third item is used to transfer the knowledge $Z_s$ of the source domain to the target domain. When $\eta = 0$, DSLR is DSLR.

In short, TDLSR summarizes DLSR from the perspective of transfer learning, but it has more transfer learning capabilities than DLSR and has better applicability. Similar to DLSR, the objective function of TDLSR can also be solved using an alternate optimization strategy. The specific derivation process is as follows:

(1) Fix $W$, and update $Z$ and $p$. Let $L = Y + B\Theta W$; formula (15) can be rewritten as

$$J(Z, p) = \arg \min_{Z,p} \left\| XZ + ep^{\mathrm{T}} - L \right\|_{\mathrm{F}}^2 + \lambda \|Z\|_{\mathrm{F}}^2 \tag{16}$$
$$+ \eta \|Z - Z_S\|_{\mathrm{F}}^2.$$

According to the optimization theory, we take a partial derivative of $p$, namely

$$\frac{\partial J(Z, p)}{\partial p} = 0,$$
$$\Longrightarrow Z^{\mathrm{T}} X^{\mathrm{T}} e + pe^{\mathrm{T}} e - L^{\mathrm{T}} e = 0, \tag{17}$$
$$\Longrightarrow p = \frac{\left(L^{\mathrm{T}} e - Z^{\mathrm{T}} X^{\mathrm{T}} e\right)}{N}.$$

Furthermore, we find the partial derivative of $Z$, we can get

$$\frac{\partial J(Z)}{\partial Z} = 0,$$
$$\Longrightarrow 2X^{\mathrm{T}} XZ - 2X^{\mathrm{T}} \left(ep^{\mathrm{T}} - L\right) + 2\lambda Z + 2\eta(Z - Z_S) = 0,$$
$$\Longrightarrow X^{\mathrm{T}} XZ - X^{\mathrm{T}} \left(e \frac{\left(e^{\mathrm{T}} L - e^{\mathrm{T}} XZ\right)}{N} - L\right) + \lambda Z$$
$$+ \eta(Z - Z_S) = 0,$$
$$\Longrightarrow X^{\mathrm{T}} \left(I_N - \frac{ee^{\mathrm{T}}}{N}\right) XZ - X^{\mathrm{T}} \left(I_N - \frac{ee^{\mathrm{T}}}{N}\right) L$$
$$+ (\lambda + \eta)Z - \eta Z_S = 0,$$
$$\Longrightarrow Z = \left(X^{\mathrm{T}} HX + (\lambda + \eta)I_d\right)^{-1} \left(X^{\mathrm{T}} HL + \eta Z_S\right)$$
$$\tag{18}$$

(2) Fix $Z$ and $p$, and update $W$. Let $G = XZ + ep^T - Y$; then, $W$ can be solved as:

$$\arg \min_{W} \|G - B\Theta W\|_{\mathrm{F}}^2.$$
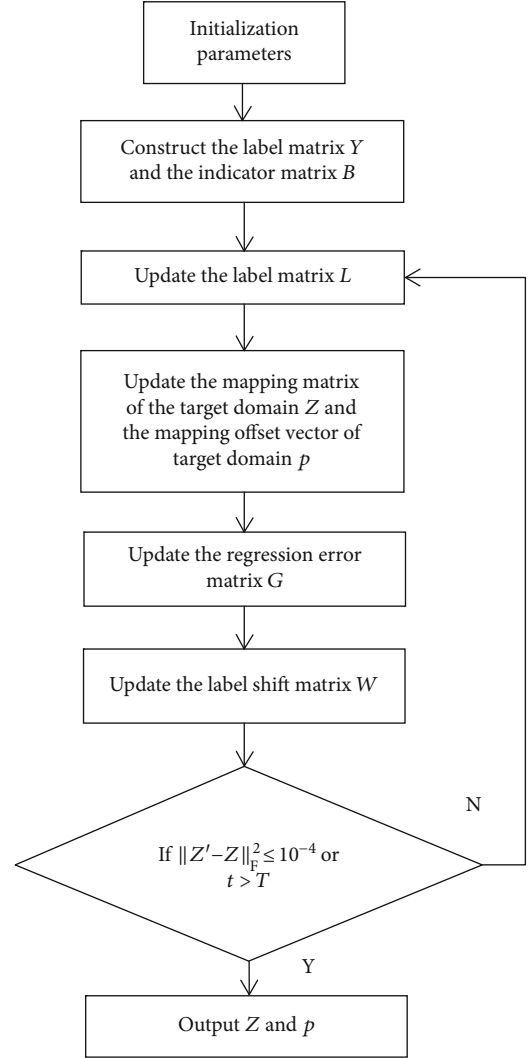$$s.t. W \geq 0 \tag{19}$$



Figure 5: The process of the TDLSR algorithm.

Similar to the way of solving $W$ in DLSR, the final solution formula of $W$ is:

$$W = \max (B\Theta G, 0). \tag{20}$$

According to the above derivation process, Algorithm 2 gives a detailed description of the TDLSR algorithm.

To understand the TDLSR algorithm more clearly, Figure 5 shows the specific process of the TDLSR.

## 4. Experimental Research

*4.1. The Environment and Parameter Settings.* To evaluate the performance of the TDLSR, a lot of experiments were carried out based on EEG signal recognition. Experiments will be conducted from two aspects: (1) the comparison experiments with classic LSR, SRC, RLR, DLSR, and SVM and (2) the comparison experiments with related methods with transform ability, such as AuSVM, Tr-Adaboost,

TABLE 2: The parameter settings in the experiments.

| Methods | Parameter settings |
|---|---|
| LSR | The regularization parameter $\lambda \in \left\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10, 10^{2}\right\}$ |
| SRC | The regularization parameter $\alpha \in \left\{10^{-3}, 5 \times 10^{-3}, 10^{-2}, 2 \times 10^{-2}, 5 \times 10^{-2}, 10^{-1}\right\}$ |
| RLR | The regularization parameter $\lambda \in \left\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}\right\}$; the heart parameter $\delta \in \left\{2^{-5}, 2^{-4}, 2^{-3}, \cdots, 2^{3}, 2^{4}, 2^{5}\right\}$ |
| DLSR | The regularization parameter $\lambda \in \left\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10, 10^{2}\right\}$ |
| SVM | The penalty factor $c \in \left\{2^{-5}, 2^{-4}, 2^{-3}, \cdots, 2^{3}, 2^{4}, 2^{5}\right\}$, the kernel parameter $\delta \in \left\{2^{-5}, 2^{-4}, 2^{-3}, \cdots, 2^{3}, 2^{4}, 2^{5}\right\}$ |
| TSVM | The Lagrange multiplier upper bound $c \in \left\{2^{-5}, 2^{-4}, 2^{-3}, \cdots, 2^{3}, 2^{4}, 2^{5}\right\}$, the kernel parameter $\delta \in \left\{2^{-5}, 2^{-4}, 2^{-3}, \cdots, 2^{3}, 2^{4}, 2^{5}\right\}$ |
| Au-SVM | The penalty factor $c \in \left\{2^{-5}, 2^{-4}, 2^{-3}, \cdots, 2^{3}, 2^{4}, 2^{5}\right\}$, the kernel parameter $\delta \in \left\{2^{-5}, 2^{-4}, 2^{-3}, \cdots, 2^{3}, 2^{4}, 2^{5}\right\}$ |
| Tr-Adaboost | The penalty factor $c \in \left\{2^{-5}, 2^{-4}, 2^{-3}, \cdots, 2^{3}, 2^{4}, 2^{5}\right\}$, the kernel parameter $\delta \in \left\{2^{-5}, 2^{-4}, 2^{-3}, \cdots, 2^{3}, 2^{4}, 2^{5}\right\}$ |
| LMPROJ | The regularization parameters $\tau \in \{0.25, 0.1, \cdots, 200\}$, $\sigma \in \{0.25, 0.1, \cdots, 200\}$, the kernel parameter $\delta \in \{0.25, 0.1, \cdots, 200\}$ |
| TDLSR | The regularization parameter $\lambda \in \left\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10, 10^{2}\right\}$, $\eta \in \left\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10, 10^{2}\right\}$ |

TABLE 3: The construction of experimental datasets.

| The subdataset | Source domain | Target domain |
|---|---|---|
| D1 | AE-each 75 | AE-each 25 |
| D2 | BDE-each 75 | BDE-each 25 |
| D3 | ABCD-each 75 | ABCD-each 25 |
| D4 | BCDE-each 75 | BCDE-each 25 |
| D5 | BE-each 75 | BC-each 25 |
| D6 | ACE-each 75 | BCE-each 25 |
| D7 | ADE-each 75 | BDE-each 25 |
| D8 | ACDE-each 75 | BCDE-each 25 |

TABLE 4: Experimental results based on nontransfer learning algorithms.

| Datasets | LSR | SRC | RLR | DLSR | SVM |
|---|---|---|---|---|---|
| D1 | 0.87 | 0.84 | 0.89 | 0.90 | 0.82 |
| D2 | 0.79 | 0.72 | 0.81 | 0.80 | 0.74 |
| D3 | 0.71 | 0.62 | 0.72 | 0.72 | 0.68 |
| D4 | 0.71 | 0.64 | 0.70 | 0.72 | 0.66 |
| D5 | 0.82 | 0.78 | 0.83 | 0.84 | 0.78 |
| D6 | 0.73 | 0.60 | 0.72 | 0.74 | 0.61 |
| D7 | 0.70 | 0.60 | 0.70 | 0.72 | 0.59 |
| D8 | 0.63 | 0.53 | 0.65 | 0.64 | 0.50 |

TSVM, and LMPROJ. The hardware environment used in the experiment is Intel (R) Core(TM) I7-9700 3.0 GHz×2, 8 G RAM, and the software environment is Windows 10 64 bit, MATLAB R2012b. Table 2 shows the parameter settings of the above algorithms.

*4.2. Experimental Dataset.* In the experiments, the EEG dataset used for epilepsy is the Bonn dataset [38, 39], which was collected by Andrzejak et al. at an epilepsy center at the University of Bonn. The EEG dataset contains five datasets, denoted by A to E. This dataset compares the EEG of the patient during the onset and nononset period with the EEG of the normal person. Dataset A and dataset B are EEG signals collected by healthy testers with their eyes open and closed. Dataset C and dataset D are the EEG signals collected by epilepsy patients outside and inside the lesion during the seizure period, and dataset E is the EEG signals collected by the patients in dataset C and dataset D during the seizure. Each of the 5 datasets includes 100 single-channel EEGs (that is 100 samples), the sampling frequency is 173.61 Hz, each segment of the signal collects 4097 frequency points, and each EEG segment lasts 23.6 s.

In our experiments, the use of the Bonn dataset is significantly different from many previous works. We constructed 8 subdatasets from the original 5 datasets to simulate different scenarios in the experiments. The source and target

TABLE 5: Experimental results based on transfer learning algorithms.

| Datasets | Au-SVM | TSVM | Tr-Adaboost | LMPROJ | TDLSR |
|---|---|---|---|---|---|
| D1 | 0.90 | 0.92 | 0.89 | 0.92 | 0.96 |
| D2 | 0.82 | 0.85 | 0.82 | 0.84 | 0.90 |
| D3 | 0.72 | 0.75 | 0.73 | 0.76 | 0.83 |
| D4 | 0.72 | 0.75 | 0.71 | 0.78 | 0.82 |
| D5 | 0.85 | 0.86 | 0.84 | 0.87 | 0.91 |
| D6 | 0.75 | 0.78 | 0.75 | 0.80 | 0.83 |
| D7 | 0.71 | 0.78 | 0.74 | 0.78 | 0.83 |
| D8 | 0.68 | 0.70 | 0.66 | 0.70 | 0.73 |

domains of the 8 subdatasets of experiments are composed of the partial data extracted from the 5 sets. We randomly select 75% of the data from a certain dataset as the source domain, and the remaining 25% as the target domain. The sample data of source and target domains of subdatasets 1-4 are derived from the same distribution, but the samples taken are different. The sample data of the source domain and the target domain of subdatasets 5-8 have different distributions. Finally, for the data in the target domain, 20% is randomly selected for testing, and the remaining 80% is used for
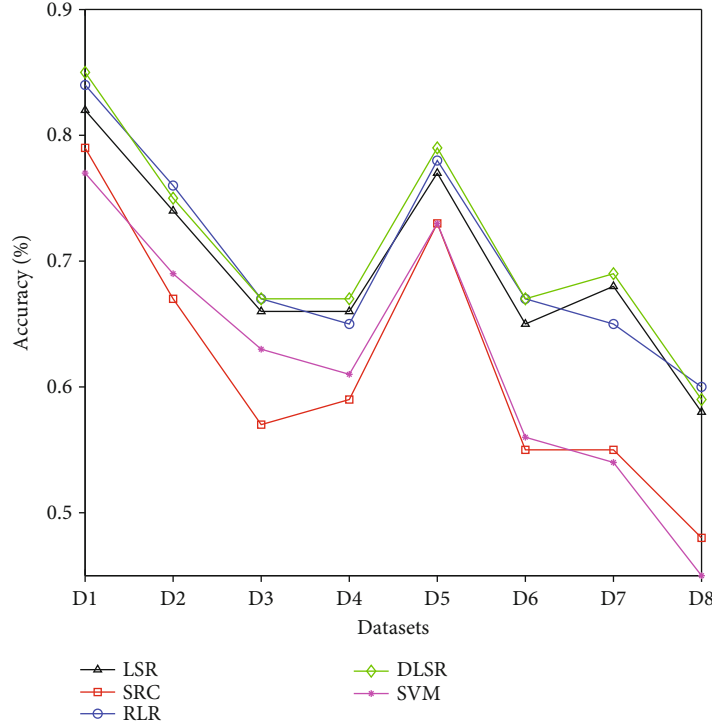
FIGURE 6: The average classification accuracy of all nontransfer learning algorithms.

training. In addition, the training and test datasets do not contain the same samples and are independent of each other. Table 3 lists the detailed information of the 8 subdatasets.

*4.3. Experimental Results and Analysis.* To verify the effectiveness of TDLSR in epilepsy EEG data recognition, the comparative experiments were conducted among the classic algorithms. The experimental results are shown in Tables 4 and 5, respectively.

As shown in the above experimental results, the conclusions are summarized as follows:

(1) In the case of the same distribution of the source domain dataset and the target domain dataset, both the nontransfer learning algorithms and the transfer learning algorithms achieve good classification results. For the datasets with certain differences in distribution, the classification effects of algorithms without transforming abilities are quite different, and the experimental results of the algorithms in Table 5 are generally better than results of the algorithms in Table 4

(2) As a whole, the performance of TDLSR introduced in this paper is obviously superior to all other algorithms. It means that by using the knowledge transferred from the source domain to the target domain, the TDLSR algorithm obtains better performance and becomes effective for epilepsy EEG recognition

(3) Comparing the performance of the algorithms in Table 4, it can be found that DLSR has the best performance, while SRC and SVM have poor performance. This is because DLSR expands the ability to distinguish between classes by using the class information in the label space for the classification task. As the results shown in Table 5, the TDLSR algorithm performs best. This is because it not only inherits the advantages of DLSR by increasing the interval between different classes but also transfers more useful information from the source domain to the target domain. It has stronger transfer learning ability than several other algorithms. And because the TDLSR method requires fewer parameters to adjust, it is easier to use, and the stability and fault tolerance are stronger than the transfer learning algorithms such as LMPROJ.

In addition, to further observe and compare the overall classification performance of all algorithms, Figures 6 and 7 also give the average classification accuracy of each algorithm on all datasets.

Observing Figures 6 and 7, it can be seen that the classification accuracy of the algorithms on the same distributed dataset is higher than the result on the different distributed dataset. Secondly, compared with the traditional nontransfer learning algorithms, the algorithm with transfer learning ability has more advantages in classification performance, and the TDLSR has a significant improvement in classification performance. Finally, for all datasets with different classes, the performance of all algorithms decreases continuously
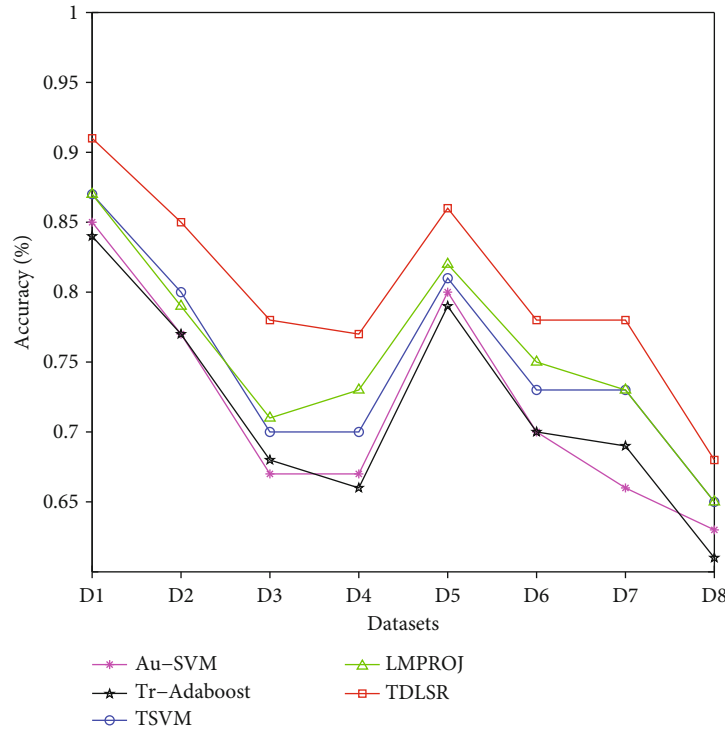
FIGURE 7: The average classification accuracy of all transfer learning algorithms.

as the number of categories increases. This is because as the number of classes increases, the information in the label space becomes more complicated, and learning the information in the label space becomes more difficult.

Meanwhile, to further verify the reliability and stability of the algorithms, we randomly added 15% white Gaussian noise to the data in the source domain to prove that the algorithm in this paper can adapt to more complex scenarios. The experimental results are shown in Tables 6 and 7, respectively.

Through the above experimental results, it can be found that the classification performance of all nontransfer learning algorithms under noisy conditions decreases more. The reason is that they cannot obtain useful knowledge from noisy data (source domain) for classification. However, all the algorithms in Table 7 can transfer some useful knowledge from the source domain for classification in the target domain, so their performance is better than the results in Table 6.

In summary, the TDLSR algorithm introduced in the paper is superior to the other algorithms in the detection of epilepsy EEG signals. And, it is easy to learn and train, has high stability, and shows certain advantages compared with other intelligent algorithms.

## 5. Conclusion

To solve the problem of serious shortage of training data in the current scene and improve the accuracy of classification, a novel DLSR-based inductive transfer learning algorithm (TDLSR) was introduced for the detection of epilepsy EEG signals. It can take advantage of both inductive transfer learn-

TABLE 6: Experimental results based on nontransfer learning algorithms with 15% noise in the source domain.

| Datasets | LSR | SRC | RLR | DLSR | SVM |
|---|---|---|---|---|---|
| D1 | 0.83 | 0.80 | 0.84 | 0.85 | 0.77 |
| D2 | 0.74 | 0.67 | 0.76 | 0.76 | 0.70 |
| D3 | 0.66 | 0.59 | 0.68 | 0.69 | 0.65 |
| D4 | 0.66 | 0.61 | 0.67 | 0.68 | 0.63 |
| D5 | 0.78 | 0.74 | 0.80 | 0.80 | 0.74 |
| D6 | 0.65 | 0.56 | 0.69 | 0.69 | 0.58 |
| D7 | 0.64 | 0.57 | 0.67 | 0.71 | 0.56 |
| D8 | 0.60 | 0.50 | 0.62 | 0.61 | 0.47 |

TABLE 7: Experimental results based on transfer learning algorithms with 15% noise in the source domain.

| Datasets | Au-SVM | TSVM | Tr-Adaboost | LMPROJ | TDLSR |
|---|---|---|---|---|---|
| D1 | 0.87 | 0.86 | 0.89 | 0.89 | 0.93 |
| D2 | 0.79 | 0.79 | 0.82 | 0.81 | 0.87 |
| D3 | 0.69 | 0.70 | 0.72 | 0.73 | 0.80 |
| D4 | 0.69 | 0.68 | 0.72 | 0.75 | 0.79 |
| D5 | 0.82 | 0.81 | 0.83 | 0.84 | 0.88 |
| D6 | 0.72 | 0.72 | 0.75 | 0.77 | 0.80 |
| D7 | 0.68 | 0.71 | 0.75 | 0.75 | 0.80 |
| D8 | 0.65 | 0.63 | 0.67 | 0.67 | 0.70 |

ing and DLSR. On the one hand, it can not only protect the security of the source domain data but also use the data of the target domain and the knowledge of the source domain

to get better performance. On the other hand, it inherits the DLSR's characteristics of being more suitable for real classification scenarios by expanding the interval between different categories. Therefore, compared with DLSR, the new algorithm not only enhances the ability of transfer learning but also ensures that the model is more reasonable. The results reflect that the improved algorithm has more advantages in epilepsy EEG signal classification compared with the traditional algorithms. However, it is found that the results is easily affected by the parameters. In a word, the quality of the parameter selection will directly affect the final detection accuracy. Therefore, to obtain higher detection accuracy, it is worth to further study the characteristics of the EEG signal to guide the setting range of the parameters in the transfer learning algorithm.

## Data Availability

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

## Conflicts of Interest

The authors declare no conflict of interest regarding the publication.

## References

[1] K. S. Wilcox, P. J. West, and C. S. Metcalf, "The current approach of the Epilepsy Therapy Screening Program contract site for identifying improved therapies for the treatment of pharmacoresistant seizures in epilepsy," *Neuropharmacology*, vol. 166, p. 107811, 2020.

[2] L. Harris and H. Angus-Leppan, "Epilepsy: diagnosis, classification and management," *Medicine*, 2020.

[3] T. Gandhi, B. K. Panigrahi, M. Bhatia, and S. Anand, "Expert model for detection of epileptic activity in EEG signature," *Expert System with Applications*, vol. 37, no. 4, pp. 3513–3520, 2010.

[4] R. Sahu, S. R. Dash, L. A. Cacha, R. R. Poznanski, and S. Parida, "Epileptic seizure detection: a comparative study between deep and traditional machine learning techniques," *Journal of Integrative Neuroscience*, vol. 19, no. 1, pp. 1–9, 2020.

[5] S. Panahi, T. Shirzadian, M. Jalili, and S. Jafari, "A new chaotic network model for epilepsy," *Applied Mathematics and Computation*, vol. 346, no. 1, pp. 395–407, 2019.

[6] S. Panahi, Z. Aram, S. Jafari, J. Ma, and J. C. Sprott, "Modeling of epilepsy based on chaotic artificial neural network," *Chaos, Solitons & Fractals*, vol. 105, pp. 150–156, 2017.

[7] S. T. George, M. S. P. Subathra, N. J. Sairamya, L. Susmitha, and M. J. Premkumar, "Classification of epileptic EEG signals using PSO based artificial neural network and tunable-Q wavelet transform," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 2, pp. 709–728, 2020.

[8] S. P. Kumar, N. Sriraam, P. G. Benakop, and B. C. Jinaga, "Entropies based detection of epileptic seizures with artificial neural network classifiers," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3284–3291, 2010.

[9] E. D. Uneyli, "Lyapunov exponents/probabilistic neural networks for analysis of EEG signals," *Expert Systems with Applications*, vol. 37, no. 2, pp. 985–992, 2010.

[10] M. Sharma, S. Patel, and U. R. Acharya, "Automated detection of abnormal EEG signals using localized wavelet filter banks," *Pattern Recognition Letters*, vol. 133, pp. 188–194, 2020.

[11] S. Raghu, N. Sriraam, Y. Temel, S. V. Rao, A. S. Hegde, and P. L. Kubben, "Performance evaluation of DWT based sigmoid entropy in time and frequency domains for automated detection of epileptic seizures using SVM classifier," *Computers in Biology and Medicine*, vol. 110, pp. 127–143, 2019.

[12] A. Temko, E. Thomas, W. Marnane, G. Lightbody, and G. Boylan, "EEG-based neonatal seizure detection with support vector machines," *Clinical Neurophysiology*, vol. 112, no. 3, pp. 464–473, 2011.

[13] N. Mahmoodian, A. Boese, M. Friebe, and J. Haddadnia, "Epileptic seizure detection using cross-bispectrum of electroencephalogram signal," *Seizure: European Journal of Epilepsy*, vol. 66, pp. 4–11, 2019.

[14] S. Ramakrishnan and A. S. Muthanantha Murugavel, "Epileptic seizure detection using fuzzy-rules-based sub-band specific features and layered multi-class SVM," *Pattern Analysis and Applications*, vol. 22, no. 3, pp. 1161–1176, 2019.

[15] H. S. Alaei, M. A. Khalilzadeh, and A. Gorji, "Optimal selection of SOP and SPH using fuzzy inference system for on-line epileptic seizure prediction based on EEG phase synchronization," *Australasian Physical & Engineering Sciences in Medicine*, vol. 42, no. 4, pp. 1049–1068, 2019.

[16] J. M. Lichtenberg and O. Simsek, "Simple regression models," in *Imperfect Decision Makes: Admitting Real-Word Rationality*, pp. 13–25, Barcelona, 2017.

[17] N. Imran, R. Tongeri, and M. Bennamoun, "Linear regression for face classification," *IEEE Transactions on Patter Analysis and Machine Intelligence*, vol. 32, no. 11, pp. 2106–2112, 2011.

[18] B. Litt, R. Esteller, J. Echauz et al., "Epileptic seizures may begin hours in advance of clinical onset," *Neuron*, vol. 30, no. 1, pp. 51–64, 2001.

[19] A. Sharmila and P. Geethanjali, "Effect of filtering with time domain features for the detection of epileptic seizure from EEG signals," *Journal of medical engineering & technology*, vol. 42, no. 3, pp. 217–227, 2018.

[20] J. Gotman, "Automatic seizure detection: improvements and evaluation," *Electroencephalography and Clinical Neurophysiology*, vol. 76, no. 4, pp. 317–324, 1990.

[21] A. Sharmila and P. Mahalakshmi, "Wavelet-based feature extraction for classification of epileptic seizure EEG signal," *Journal of Medical Engineering &amp; Technology*, vol. 41, no. 8, pp. 670–680, 2017.

[22] I. Goker, O. Osman, S. Ozekes, M. B. Baslo, M. Ertas, and Y. Ulgen, "Classification of juvenile myoclonic epilepsy data acquired through scanning electromyography with machine learning algorithms," *Journal of Medical Systems*, vol. 36, no. 5, pp. 2705–2711, 2012.

[23] O. Faust, U. R. Acharya, L. C. Min, and B. H. Sputh, "Automatic identification of epileptic and background EEG signals using frequency domain parameters," *International journal of neural systems*, vol. 20, no. 2, pp. 159–176, 2012.

[24] R. Hopfengartner, F. Kerling, V. Bauer, and H. Stefan, "An efficient, robust and fast method for the offline detection of epileptic seizures in long-term scalp EEG recordings," *Clinical Neurophysiology*, vol. 118, no. 11, pp. 2332–2343, 2007.

[25] A. Maheshwari, "Rodent EEG: expanding the spectrum of analysis," *Epilepsy Currents*, vol. 20, no. 3, pp. 149–153, 2020.

[26] S. Ashok and G. Purusothaman, "A review on the pattern detection methods for epilepsy seizure detection from EEG signals," *Biomedical Engineering/Biomedizinische Technik*, vol. 64, no. 5, pp. 507–517, 2019.

[27] R. B. Pachori, "Analysis of normal and epileptic seizure EEG signals using empirical mode decomposition," *Computer Methods and Programs in Biomedicine*, vol. 104, no. 3, pp. 373–381, 2011.

[28] R. J. Martis, U. R. Acharya, H. J. Tan et al., "Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals," *International Journal of Neural Systems*, vol. 22, no. 6, 2012.

[29] V. Bajaj and R. B. Pachri, "Epileptic seizure detection based on the instantaneous area of analytic intrinsic mode functions of EEG signals," *Biomedical Engineering Letters*, vol. 3, no. 1, pp. 17–21, 2013.

[30] R. B. Pachori, R. Sharma, and S. Patidar, "Classification of normal and epileptic seizure EEG signals based on empirical mode decomposition," in *Complex system modelling and control through intelligent soft computations*, pp. 367–388, Springer, Cham, 2015.

[31] C. Guerrero-Mosquera, A. M. Trigueros, J. I. Franco, and Á. Navia-Vázquez, "New feature extraction approach for epileptic EEG signal detection using time-frequency distributions," *Medical &Biological Engineering & Computing*, vol. 48, no. 4, pp. 321–330, 2010.

[32] E. Ebrahimzadeh, S. M. Alavi, A. Bijar, and A. Pakkhesal, "A novel approach for detection of deception using Smoothed Pseudo Wigner-Ville Distribution (SPWVD)," *Journal of Biomedical Science and Engineering*, vol. 6, no. 1, pp. 8–18, 2013.

[33] H. Polat, M. U. Aluçlu, and M. S. Özerdem, "Evaluation of potential auras in generalized epilepsy from EEG signals using deep convolutional neural networks and time-frequency representation," *Biomedical Engineering / Biomedizinische Technik*, 2019.

[34] M. A. Al-Manie and W. J. Wang, "A modified S-transform for EEG signals analysis," *Internation Journal of Computational Methods*, vol. 12, no. 3, 2015.

[35] N. Kannathal, M. L. Choo, U. R. Acharya, and P. K. Sadasivan, "Entropies for detection of epilepsy in EEG," *Computer Methods and Programs in Biomedicine*, vol. 80, no. 3, pp. 187–194, 2005.

[36] A. Arnold, R. Nallapati, and W. W. Cohen, "A comparative study of methods for transductive transfer learning," in *Proceedings of the 7th IEEE International Conference on Data Mining Workshops(ICDMW'07)*, pp. 77–82, IEEE Computer Society, Washington, DC, USA, 2007.

[37] S. Xiang, F. Nie, G. Meng, C. Pan, and C. Zhang, "Discriminative least squares regression for multiclass classification and feature selection," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 11, pp. 1738–1754, 2012.

[38] Y. Jiang, Z. Deng, F.-L. Chung et al., "Recognition of epileptic EEG signals using a novel multiview TSK fuzzy system," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 1, pp. 3–20, 2017.

[39] Y. Jiang, D. Wu, Z. Deng et al., "Seizure classification from EEG signals using transfer learning, semi-supervised learning and TSK fuzzy system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 12, pp. 2270–2284, 2017.