Enquiries:
If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines)

WILEY | Hindawi

Research Article

Identification of Alcoholism Based on Wavelet Renyi Entropy and Three-Segment Encoded Jaya Algorithm

Shui-Hua Wang,[1,2,3] **Khan Muhammad** (ID)**,**[4] **Yiding Lv,**[5] **Yuxiu Sui** (ID)**,**[5]
Liangxiu Han (ID)**,**[6] **and Yu-Dong Zhang** (ID)[1,2,3]

[1]*School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, Henan 454000, China*
[2]*Department of Informatics, University of Leicester, Leicester LE1 7RH, UK*
[3]*School of Computer Science and Technology, Nanjing Normal University, Nanjing, Jiangsu 210023, China*
[4]*Digital Contents Research Institute, Sejong University, Seoul, Republic of Korea*
[5]*Department of Psychiatry, Nanjing Brain Hospital, Nanjing Medical University, Nanjing, Jiangsu 210029, China*
[6]*School of Computing, Mathematics and Digital Technology (SCMDT), Manchester Metropolitan University,*
 Manchester M156BH, UK

Correspondence should be addressed to Khan Muhammad; khanmuhammad@sju.ac.kr, Yuxiu Sui; suiyuxiu@aliyun.com,
Liangxiu Han; l.han@mmu.ac.uk, and Yu-Dong Zhang; yudongzhang@ieee.org

The alcohol use disorder (AUD) is an important brain disease, which could cause the damage and alteration of brain structure. The current diagnosis of AUD is mainly done manually by radiologists. This study proposes a novel computer-vision-based method for automatic detection of AUD based on wavelet Renyi entropy and three-segment encoded Jaya algorithm from MRI scans. The wavelet Renyi entropy is proposed to provide multiresolution and multiscale analysis of features, describe the complexity of the brain structure, and extract the distinctive features. Grid search method was used to select the optimal wavelet decomposition level and Renyi order. The classifier was constructed based on feedforward neural network and a three-segment encoded (TSE) Jaya algorithm providing parameter-free training of the weights, biases, and number of hidden neurons. We have conducted the experimental evaluation on 235 subjects (114 are AUDs and 121 healthy). k-fold cross validation has been used to avoid overfitting and report out-of-sample errors. The results showed that the proposed method outperforms four state-of-the-art approaches in terms of accuracy. The proposed TSE-Jaya provides a better performance, compared to the conventional approaches including plain Jaya, multiobjective genetic algorithm, particle swarm optimization, bee colony optimization, modified ant colony system, and real-coded biogeography-based optimization.

1. Introduction

Alcohol use disorder (AUD) affected 208 million people worldwide in 2010. It can cause severe adverse effects to the brain, liver, heart, and pancreas. The long-term misuse can lead to increased tolerance to alcohol, making it difficult to control the consumption. The short-term misuse can lead to "blood alcohol concentration (BAC)." A BAC from 0.35% to 0.80% can cause fatal respiratory depression and life-threating alcohol poisoning.

This paper studies the effect of long-term alcohol misuse on the brain. The alcohol misuse can have a damaging effect on the brain neurons; hence, patients with long-term

AUD have smaller volumes of white matter and gray matter than age-matched controls. Besides, alcohol causes adverse effect on the prefrontal cortex and cerebellum. The current diagnosis of AUD mainly relies on manual observation based on brain images. However, due to mild symptoms, the radiologists may miss the slight shrinkage of AUD brains and be unable to identify it at an early stage. It is necessary to create an efficient approach that can monitor the patient brain via magnetic resonance imaging (MRI) and provide automatic, early diagnosis.

Over the last decades, computer-vision-based techniques have been proposed for automatically detecting changes on brain structure for brain related disease diagnosis based on

TABLE 1: Demographic characteristics.

	Alcoholic		Nonalcoholic	
	Women (56)	Men (58)	Women (62)	Men (59)
Age (y)	59.0 + 8.0	56.5 + 8.9	56.9 + 8.4	55.3 + 7.9
Education (y)	9.9 + 1.9	9.3 + 1.6	9.6 + 2.4	9.3 + 2.5
DHD (y)	13.2 + 3.5	19.6 + 5.4	0.0 ± 0.0	0.0 ± 0.0
DDE (grams/d)	197.6 + 62.1	300.7 + 92.1	6.8 + 4.7	4.9 + 3.8
LOS (y)	9.8 + 5.1	6.7 + 3.4	-	-
AUDIT	25.1 + 4.2	25.4 + 4.9	1.5 + 2.0	1.6 + 2.2

(DDE, daily drinks of ethanol; DHD, duration of heavy drinking; AUDIT, alcohol use disorders identification test; LOS, length of sobriety).

MRI scans. Nayak et al. (2016) [1] presented a brain image classification algorithm based on random forest. Alweshah and Abdullah (2015) [2] hybridized firefly algorithm (FA) and probabilistic neural network (PNN). They used the proposed method for detecting changes in the brain. Lv and Hou [3] proposed an improved particle swarm optimization (IPSO) to detect alcoholism in MRI scanning. Monnig (2012) [4] suggested detecting white matter atrophy in neuroimaging of AUD. Yang (2017) [5] combined Hu moment invariant (HMI) and support vector machine (SVM) for pathological brain detection. Jiang and Zhu (2017) [6] explored the method using pseudo Zernike moment (PZM). Lv and Sui (2017) [7] used data augmentation technique for alcoholism detection.

Although several of the above methods were developed for pathological brain detection, they can be easily transferred and applied to alcoholism detection. Nevertheless, these methods suffer from several common problems: first, approaches that do not take into account expressions of complexity of the brain structure do not exhibit good performance in AUD. Second, the training algorithms of existing classifiers may fall into local optimal and it is difficult to optimize the hyperparameters (e.g., the number of hidden neurons in a feedforward neural network) of the classifiers [8].

To address the above problems, we propose in this study a novel identification method of alcoholic use disorder. Our contributions include the following: (1) a novel feature extraction method—wavelet Renyi entropy, which can describe the complexity of brain structure at multiple scales—and (2) an improved Jaya algorithm to train a feedforward neural network, which can optimize the weights, biases, and the number of hidden neurons simultaneously. Our training algorithm does not need to set algorithm-specific parameters.

The rest of this paper is organized as follows: Section 2 describes the subjects, scan protocol, and slice selection method. Section 3 presents the proposed feature extraction method—wavelet Renyi entropy. Section 4 describes the classifier construction and the proposed training algorithm: three-segment encoded Jaya algorithm. Section 5 provides the implementation procedure and the evaluation method. Besides, we show how to use grid search to optimize the parameters of wavelet Renyi entropy. The results and discussions are presented in Section 6. Section 7 concludes the work.

2. Materials

2.1. Subjects. The subjects went through a medical history interview to guarantee they met the inclusion criteria. Those qualified applicants received the computerized diagnostic interview schedule version IV, which ascertains the presence or absence of major psychiatric disorders. Applicants were excluded if mandarin was not their first language, if they were left-handed, or if they had HIV, epilepsy, and stroke; Wernicke–Korsakoff syndrome; bipolar disorder; cirrhosis or liver failure or seizures unrelated to alcoholism, head injury with loss of consciousness more than 15 minutes unrelated to alcoholism, depression, schizophrenia, and other psychotic disorders.

Finally, we enrolled 114 abstinent long-term chronic alcoholic participants (58 men and 56 women) and 121 nonalcoholic control participants (59 men and 62 women). Participants were enrolled through flyers posted in Jiangsu Province Hospital, Nanjing Children's Hospital, and Nanjing Brain Hospital, as well as the Internet-based advertisements. The data collection lasted for a total of three years. The research was approved by the Institutional Review Board of the participating hospitals. Informed consent was obtained from each participant.

The 235 participants were tested by the "Alcohol Use Disorder Identification Test (AUDIT)" [9]. The unit "ounce" was transformed to "gram," since the former is not widely identified in China. Their demographic characteristics are shown in Table 1. In this study, we only focus on the structural imaging data.

2.2. Scan Protocol. All 235 subjects lied down as still as possible, with their eyes closed and remaining conscious. Scanning was implemented by a Siemens Verio Tim 3.0T MR scanner (Siemens Medical Solutions, Erlangen, Germany). In total, 216 sagittal slices covering the whole brain were acquired, using an MP-RAGE sequence. The imaging parameters were listed as follows: slice thickness = 0.8 mm, TE = 2.50 ms, TR = 2000 ms, TI = 900 ms, FA = 9°, matrix = 256 × 256, and FOV = 256 mm × 256 mm. The acquired image was 16-bit gray level depth, and we reduced it to 8-bit gray level depth, since the alcoholism alters the structure of healthy brain and it does not change the gray level of brain images. Besides, 8-bit gray level provides enough information, so it is unnecessary to use 16-bit gray level images.
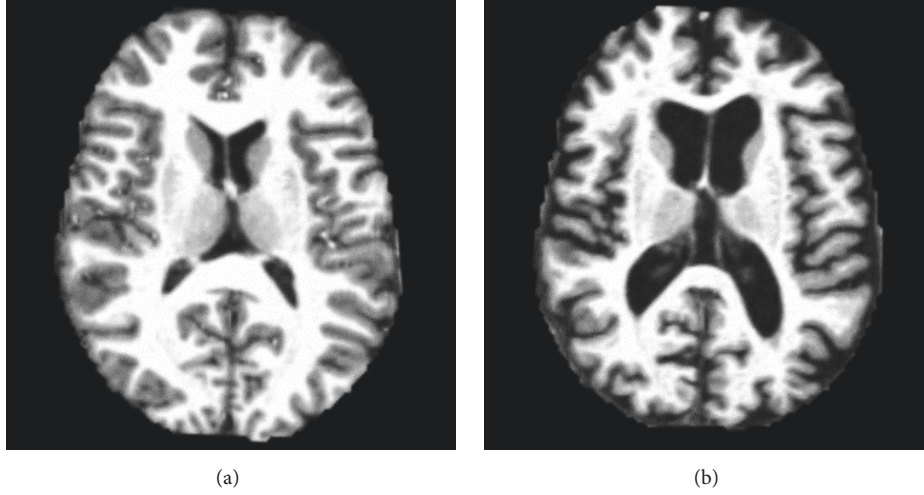
(a)

(b)

FIGURE 1: Slice examples between (a) a nonalcoholic brain and (b) an alcoholic brain.

2.3. Slice Selection. We used FMRIB software library (FSL) v5.0 software [16, 17] to extract brain and remove skulls for each scanned 3D image. All the volumetric images were normalized to a standard MNI template. Afterwards, we resampled each image to 2 mm isotropic voxel. The slice at $Z = 80$ (8 mm) at MNI_152 coordinate, which is an average of 152 T1-weighted MRI scans linearly transformed to Talairach space, was chosen for each patient. The reason for selecting the 80th slice is that it contains the two distinguishing features of alcoholic patients: (i) the enlarged ventricle and (ii) the shrunk gray matter, for example, the precentral gyrus [18], inferior frontal gyrus [19], and middle temporal gyrus [20]. Figure 1 shows the clear difference between the alcoholic and healthy samples.

Afterwards, the background was cropped, leaving a rectangle matrix with size of 176×176 for the subsequent classifier training. The datasets used in this study are available upon request.

3. The Proposed Feature Extraction Method

To extract distinctive features, this study proposed a new wavelet Renyi entropy (WRE), which combines discrete wavelet transform and Renyi entropy in order to describe the complexity of the brain structure. The wavelet decomposition provides multiresolution and multiscale analysis, while the Renyi entropy provides the complexity description of the wavelet subbands of brain structure.

3.1. Wavelet Decomposition. For a specific signal/image, the discrete wavelet transform (DWT) transforms the signal/image to the wavelet domain. It performs the transformation at multiple levels, by delivering the previous approximation subband to the quadrature mirror filters (abbreviated as QMF) [21]. Compared to traditional Fourier transform, DWT has the key advantage of temporal/spatial resolution.

Let $q(t)$ be a given one-dimensional signal, and the continuous wavelet transform of $q(t)$ is depicted as

$$O_\psi (f_s, f_t) = \int_{-\infty}^{\infty} q(t) \times \psi(t \mid f_s, f_t) \, dt, \tag{1}$$

where O represents the coefficients and $\psi(t)$ the mother wavelet. $\psi(t \mid f_s, f_t)$ is defined as

$$\psi(t \mid f_s, f_t) = \frac{1}{\sqrt{f_s}} \psi\left(\frac{t - f_t}{f_s}\right). \tag{2}$$

Here, f_s represents the scale factor and f_t the translation factor (both f_s and $f_t > 0$). Formula (1) can be discretized by replacing f_s and f_t to discrete variables a and b.

$$\begin{aligned} f_s &= 2^a \\ f_t &= b \times 2^a, \end{aligned} \tag{3}$$

where the parameters a and b represent the values of scale and translation factors, respectively. By this means, we can produce the DWT as

$$L(n \mid a, b) = \downarrow \left(\sum_n q(n) \times l_a^*(n - 2^a b) \right)$$

$$G(n \mid a, b) = \downarrow \left(\sum_n q(n) \times g_a^*(n - 2^a b) \right). \tag{4}$$

Here, n means the discrete version of variable t. $\downarrow$ means the downsampling. The functions $l(n)$ and $g(n)$ represent the low-pass filter and high-pass filter, respectively. L and G represent the approximation subband and the detail subband, respectively.

For a two-dimensional DWT (abbreviated as 2D-DWT) [22], suppose the image is symbolized as $q(m, n)$, and there are four subbands in all after each decomposition (A, H, V, and D), shown in Figure 2. The subband A is the approximation component of original image. Subbands H, V, and D represent horizontal, vertical, and diagonal position, respectively. A will be decomposed into four new subbands at a higher level, to produce corresponding higher-level subbands.

```
Step 1    Input a given brain image X.
Step 2    Perform a k-level wavelet decomposition, and obtained (3k + 1) subbands.
Step 3    For i = 1: 3k + 1
              Get the 256-bin histogram of the ith subband wavelet coefficients;
              Obtain Renyi entropy with order of α over the histogram.
          End
Step 4    Output the catenation of all Renyi entropy values of all subbands.
```

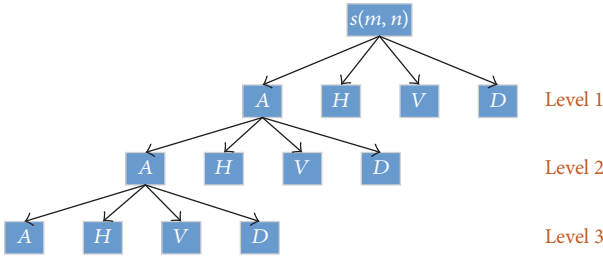PSEUDOCODE 1: Pseudocode of wavelet Renyi entropy.

FIGURE 2: Pipeline of 2D-DWT.

3.2. Renyi Entropy.

Each subband of wavelet decomposition can be regarded as a discrete variable X. Suppose X has possible outcomes as

$$X = \{x_1, x_2, \ldots, x_n\}. \tag{5}$$

Suppose the corresponding probability is defined as

$$p_i = \Pr(X = x_i). \tag{6}$$

The α-order Renyi entropy is defined as [23]

$$H_\alpha(X) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{n} p_i^\alpha\right). \tag{7}$$

The Renyi entropy is Schur-concave and it is a nonincreasing function in α. In some special cases, the Renyi entropy will turn to other types of entropies. For instance, $H_0(X)$ is called Hartley entropy, $H_1(X)$ is Shannon entropy, and $H_\infty(X)$ is the min-entropy [24].

Suppose we have a binary random variable X with $P = [p_1, p_2]$, where $p_2 = 1 - p_1$. The Renyi entropies with different α-values against p_1 are plotted in Figure 3. The concaveness and the nonincreasing against α properties are obvious from this picture. Zero values are included, since this does not affect the calculation of Renyi entropy.

3.3. Wavelet Renyi Entropy.

In the past, scholars have proposed the so-called "wavelet Renyi entropy." Nevertheless, our proposed WREs are different from traditional WREs. First, traditional WREs are mainly for one-dimensional signal, while ours are for two-dimensional image. Second, traditional WREs calculate entropies over the approximation subbands, while our method calculates entropies over both approximation and detail subbands of wavelet coefficients.
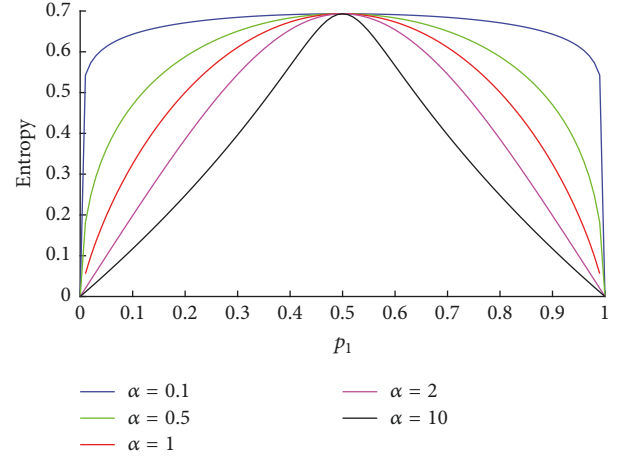
FIGURE 3: Renyi entropy with different α-values for a binary random variable.

The pseudocode of WRE is depicted in Pseudocode 1. The bin number of wavelet coefficient histogram is set to 256 in this study.

For a given image, our proposed WRE produced a $(3k + 1)$-element feature vector. Here, we choose the optimal values of α and k by grid searching approach. The detailed implementation is explained in Section 5.3.

4. The Classifier Construction Based on a Feedforward Neural Network (FNN) and Three-Segment Jaya

To train the classifier, we have proposed using a feedforward neural network (FNN) and three-segment Jaya algorithm. Scholars have used various classifiers in medical brain image analysis, such as decision tree, support vector machine [25], and naive Bayesian classifier. Nevertheless, the feedforward neural network (FNN) won remarkable success, because of the universal approximation theorem [26], which says the following.

Suppose η is a bounded and nonconstant continuous function. Given any function f and any small number $\varepsilon > 0$, there exist an integer N, real vectors ω_i, and real constants c_i and b_i, such that we have

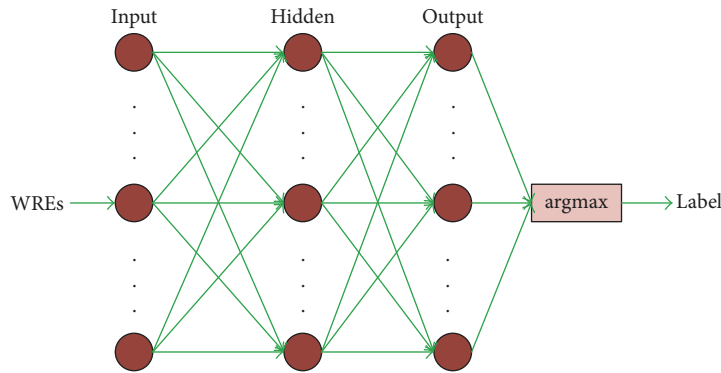$$F(x) = \sum_{i=1}^{N} c_i \eta\left(\omega_i^T x + b_i\right), \tag{8}$$

FIGURE 4: Diagram of an FNN.

where $F(x)$ can be used as an approximation realization of function f, which satisfies

$$|F(x) - f(x)| < \varepsilon. \tag{9}$$

However, the traditional FNN training algorithm is a backpropagation (BP) gradient descent algorithm. The BP and its variants often converge to local optimal points. Three-segment encoded Jaya is introduced to address the problem. In the following sections, we will detail each of the methods.

4.1. Structure of FNN. Structurally, the FNN include three layers: (i) an input layer accepted the features; (ii) a hidden layer contains hidden neurons; (iii) an output layer outputs the scores of each class. Finally, the "argmax" function predicts the class associated with the largest score. Figure 4 presents the diagram of FNN. The number of input neurons is the same as the number of features extracted from brain images, the number of output neurons is the same as the number of classes, and the number of hidden neurons is commonly obtained by hyperparameter optimization.

4.2. Jaya Algorithm. As mentioned earlier, the traditional FNN training has an issue with global optimization; to address this problem, a massive number of global optimization algorithms were proposed and employed to train FNN, particularly in the field of brain image classification. For example, Hajimani et al. (2017) [11] designed a multiobjective genetic algorithm (MOGA) to detect cerebral vascular accidents. Chen et al. (2017) [12] used particle swarm optimization (PSO) to classify MRI brain tissues. Subramaniam and Radhakrishnan (2016) [13] used bee colony optimization (BCO) to classify brain cancer image. Raghtate and Salankar (2015) [14] proposed a modified ant colony system (MACS) to realize automatic brain MRI classification. Chen and Du (2017) [15] proposed a real-coded biogeography-based optimization (RCBBO) method for pathological brain detection.

Those algorithms make the classifier more robust than BP; nevertheless, their own parameters need to be fine-tuned, which causes the hyperparameter optimization problem. To overcome the limitation of existing optimization approaches, Jaya as a powerful global optimization approach has been introduced by Rao (2016) [27] as a benchmark function for constrained and unconstrained problems. It is
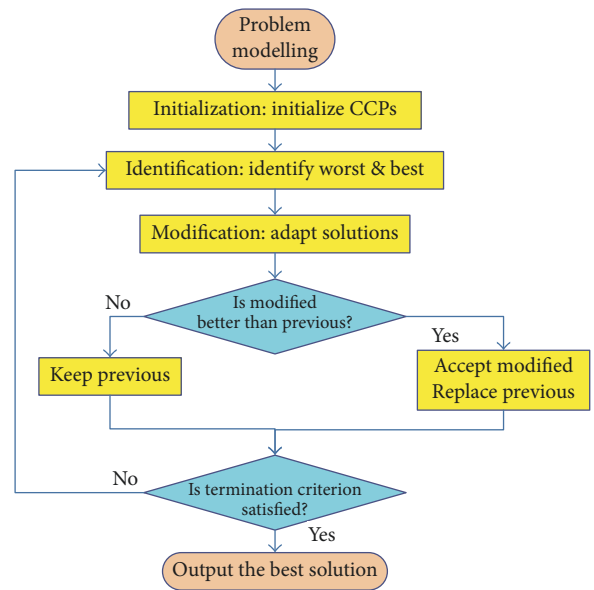
FIGURE 5: Diagram of Jaya algorithm.

an algorithm-specific parameter-free approach which has been proven to be superior to state-of-the-art optimization algorithms and has been successfully applied in thermal performance optimization [28], photovoltaic model identification [29], cooling tower design [30], sensing period adaptation [31], heat change optimization [32], and so forth.

Figure 5 shows the diagram of Jaya algorithm. Assume x, y, and z are the index of iteration, variable, and candidate. Assume $V(x, y, z)$ means the yth variable of zth solution candidate at xth step. Assume $R(x, y, 1)$ and $R(x, y, 2)$ are two positive numbers in the range of $[0, 1]$ and generated at random. The modified candidate $Q(x, y, z)$ is defined as

$$
\begin{aligned}
Q(x, y, z) = {} & V(x, y, z) + R(x, y, 1) \\
& \times [V(x, y, b) - |V(x, y, z)|] \\
& - R(x, y, 2) \\
& \times [V(x, y, a) - |V(x, y, z)|],
\end{aligned} \tag{10}
$$

where a and b denote the index of worst and best candidate within the population:

$$b = \arg\min_z \left(V(x, y, z) \right)$$
$$a = \arg\max_z \left(V(x, y, z) \right). \tag{11}$$

Hence, $V(x, y, a)$ and $V(x, y, b)$ denote the worst and best value of yth variable at xth iteration.

The 2nd term "$R(x, y, 1) \times [V(x, y, b) - |V(x, y, z)|]$" in (10) represents that the candidate needs to move closer to the best one. The 3rd term "$-R(x, y, 2) \times [V(x, y, a) - |V(x, y, z)|]$" in (10) represents that the candidate needs to move away from the worst candidate, noting the "$-$" symbol before $R(x, y, 2)$ [33]. The updated candidate at iteration $(x + 1)$ is written as

$$V(x + 1, y, z) = \begin{cases} Q(x, y, z) \\ V(x, y, z), \end{cases}$$
$$\text{if } \begin{cases} h[N(x, y, z)] < h[V(x, y, z)] \\ h[N(x, y, z)] \geq h[V(x, y, z)], \end{cases} \tag{12}$$

where h represents the fitness function.

Equation (12) indicates that $V(x + 1, y, z)$ is assigned with $Q(x, y, z)$ if the modified candidate $Q(x, y, z)$ is better in terms of fitness than $V(x, y, z)$; otherwise it is assigned with $V(x, y, z)$. The algorithm iterates until the termination criterion is satisfied. We set the termination criterion as follows: either the algorithm reaches maximum iteration epoch or the error does not reduce for five epochs.

4.3. Three-Segment Encoded Jaya. The existing Jaya algorithm is mainly used to train weights and biases of FNN as described in Phillips (2017) [10]. However, we believe the number of hidden neurons is also an important hyperparameter that influences the classification performance of FNN. Hence, we proposed a three-segment encoded Jaya algorithm (TSE-Jaya), which aims to optimize the weights, biases, and number of hidden neurons simultaneously. The candidate $V(x, y, z)$ now contains three parts as

$$V(x, y, z) = [V_1(x, y, z) \quad V_2(x, y, z) \quad V_3(x, y, z)], \tag{13}$$

where $V()_1$, $V()_2$, and $V()_3$ represent extracting the first part, second part, and third part of the solution candidate representation. The first part encodes the weights, the second part encodes the biases, and the third part encodes the number of hidden neurons (NHN). The modified candidate is defined consequently as

$$Q(x, y, z) = [Q_1(x, y, z) \quad Q_2(x, y, z) \quad Q_3(x, y, z)]. \tag{14}$$

The modification rule does not obey (10), and the new modification rule is three-fold as follows:

$$Q_1(x, y, z) = V_1(x, y, z) + R(x, y, 1)$$
$$\times [V_1(x, y, b) - |V_1(x, y, z)|]$$
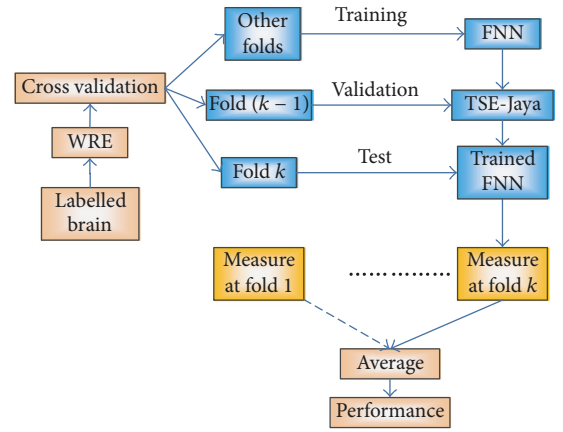

FIGURE 6: The flowchart of the proposed method.

$$- R(x, y, 2)$$
$$\times [V_1(x, y, a) - |V_1(x, y, z)|]$$
$$Q_2(x, y, z) = V_2(x, y, z) + S(x, y, 1)$$
$$\times [V_2(x, y, b) - |V_2(x, y, z)|]$$
$$- S(x, y, 2)$$
$$\times [V_2(x, y, a) - |V_2(x, y, z)|]$$
$$Q_3(x, y, z) = V_3(x, y, z) + T(x, y, 1)$$
$$\times [V_3(x, y, b) - |V_3(x, y, z)|]$$
$$- T(x, y, 2)$$
$$\times [V_3(x, y, a) - |V_3(x, y, z)|], \tag{15}$$

where S and T are two random positive numbers, similar to variable R. Other procedures are the same as those in Jaya algorithm.

5. Implementation and Evaluation

5.1. Cross Validation Based Implementation. Figure 6 presents the flowchart of our method. Here, the k-fold cross validation method [34] was used in order to avoid overfitting and report out-of-sample errors. We divide the whole dataset 10-fold. In kth trial, $(k-1)$th fold is used as validation, kth fold is used as test, and other folds are used as training. The training iterates until the accuracy over validation (a_v) set increases for five continuous epochs. For a clear understanding, we plotted a toy example in Figure 7. Here, at epoch 6, the validation error reaches the minimum. Then, from the 6th to 11th epoch, we can observe the validation increases although the training error decreases, which indicates an overfitting occurs. Hence, we should select the weights corresponding to the 6th epoch. The goal of k-fold cross validation in this study is to avoid overfitting.

TABLE 2: Grid searching of hyperparameters.

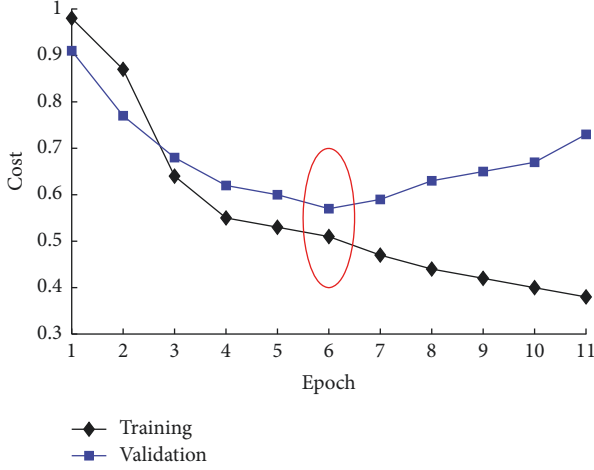| Grid searching	Implementation
Wavelet decomposition level k	Search [1, 2, 3, 4, 5] and obtain the best level k^*
Order α of Renyi entropy	*Step 1.* Search $[0, 1, 2, \ldots, 6]$ and obtain the best candidate c
	Step 2. Search $[c - 0.5, c - 0.4, \ldots, c + 0.4, c + 0.5]$ and obtain the best order α^*

FIGURE 7: A toy example explaining the effect of validation set.

After the training terminates, the measure over the test set is recorded. Finally, all the measures over all test sets of all trials are averaged, and the final classification performance was calculated. The ideal confusion matrix of one time of 10-fold cross validation is

$$I\left(f = 10, t = 1\right) = \begin{bmatrix} 114 & 0 \\ 0 & 121 \end{bmatrix}, \tag{16}$$

where I means the ideal confusion matrix, f the number of folds, and t the number of repetitions. In this study, we run the 10-fold cross validation 10 times, and the ideal confusion matrix is

$$I\left(f = 10, t = 10\right) = \begin{bmatrix} 1140 & 0 \\ 0 & 1210 \end{bmatrix}. \tag{17}$$

5.2. Evaluation. The evaluation was performed on the realistic confusion matrix of 10×10-fold cross validation. Suppose the positive class is alcoholism, and the negative class is the control. We can define true positive (TP) as alcoholism correctly identified, true negative (TN) as control correctly identified, false positive (FP) as control mispredicted as alcoholism, and false negative (FN) as alcoholism mispredicted as control. Finally, we define three measures: sensitivity (Sen), specificity (Spc), and accuracy (Acc).

$$\begin{aligned} \text{Sen} &= \frac{\text{TP}}{\text{TP} + \text{FN}} \\ \text{Spc} &= \frac{\text{TN}}{\text{TN} + \text{FP}} \\ \text{Acc} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \end{aligned} \tag{18}$$

5.3. Grid Searching. In the grid searching, the criterion uses the "*accuracy* (Acc)" measure defined above. The implementation is explained in Table 2. For wavelet decomposition level k, a simple grid searching from 1 to 5 with an increase of 1 was used, since k should be an integer.

$$k^* = \arg \max_{[1,2,\ldots,5]} \left(\text{Acc}\right). \tag{19}$$

For the order α of Renyi entropy, a coarse-to-fine searching strategy was used. First, a coarse grid was set from 0 to 6 with an increase of 1, and we obtained the coarse candidate c.

$$c = \arg \max_{[0,1,\ldots,6]} \left(\text{Acc}\right). \tag{20}$$

Then, a fine grid was set from $c - 0.5$ to $c + 0.5$ with an increase of 0.1, and the optimal order α^* is obtained as

$$\alpha^* = \arg \max_{[c-0.5, c-0.4, \ldots, c+0.5]} \left(\text{Acc}\right). \tag{21}$$

6. Results and Discussions

Our programs were developed in-house. The experiment ran on the platform of Dell laptop with 2.20 GHz Intel Core i7-4702HQ CPU and 16 GB RAM. The operating system was Windows 10. MATLAB 2017a is the programing development environment.

6.1. Statistical Analysis. The cross validation divides the dataset into 10 sets. In each run, the number of sets resulting from the division is different. We set the wavelet decomposition level as 4 and Renyi α-value as 1.2. The maximum iterative epoch is set as 1000, and the population in Jaya algorithm is set to 20. The sensitivity, specificity, and accuracy of our method are listed in Tables 3, 4, and 5, respectively. We can observe our proposed method achieved a sensitivity of $93.60 \pm 1.55\%$, a specificity of $93.72 \pm 1.42\%$, and an accuracy of $93.66 \pm 1.23\%$.

6.2. Comparison to State-of-the-Art Approaches. We compared this proposed method "WMI + FNN + TSE-Jaya" with four state-of-the-art approaches: FA + PNN [2], IPSO method [3], HMI + SVM [5], and PZM [6]. All the algorithms were run over a 10×10-fold cross validation over our dataset. The results of 10×10-fold cross validation of four state-of-the-art methods are shown in Table 6. Finally, the comparison result is presented in Table 7.

6.3. Optimal Wavelet Decomposition Level. In this experiment, we fixed the Renyi order α to 1.2 and let the wavelet decomposition level change from 1 to 5 with increase of 1. The

TABLE 3: Sensitivity of the proposed method (unit: %).

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Total
R1	100.00	90.91	100.00	81.82	100.00	83.33	100.00	100.00	100.00	90.91	94.74
R2	100.00	100.00	100.00	91.67	100.00	72.73	72.73	100.00	100.00	100.00	93.86
R3	100.00	100.00	90.91	100.00	100.00	100.00	91.67	100.00	81.82	91.67	95.61
R4	100.00	90.91	81.82	90.91	75.00	100.00	100.00	100.00	100.00	90.91	92.98
R5	100.00	81.82	100.00	100.00	100.00	100.00	63.64	90.91	90.91	91.67	92.11
R6	100.00	100.00	100.00	100.00	100.00	90.91	72.73	66.67	100.00	91.67	92.11
R7	100.00	83.33	90.91	90.91	90.91	81.82	91.67	100.00	100.00	100.00	92.98
R8	91.67	100.00	91.67	100.00	90.91	90.91	100.00	90.91	100.00	90.91	94.74
R9	90.91	100.00	100.00	100.00	100.00	90.91	90.91	91.67	100.00	91.67	95.61
R10	100.00	90.91	91.67	81.82	100.00	90.91	90.91	83.33	90.91	91.67	91.23
Avr											93.60 ± 1.55

(F, fold; R, run).

TABLE 4: Specificity of the proposed method (unit: %).

|	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Total
R1	100.00	100.00	100.00	100.00	100.00	100.00	83.33	91.67	91.67	91.67	95.87
R2	100.00	91.67	100.00	83.33	100.00	100.00	91.67	84.62	100.00	91.67	94.21
R3	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	75.00	66.67	94.21
R4	91.67	91.67	100.00	92.31	83.33	100.00	100.00	91.67	66.67	100.00	91.74
R5	91.67	83.33	100.00	91.67	91.67	92.31	100.00	83.33	100.00	91.67	92.56
R6	100.00	66.67	84.62	100.00	100.00	100.00	100.00	83.33	100.00	100.00	93.39
R7	91.67	100.00	83.33	100.00	83.33	100.00	100.00	100.00	100.00	91.67	95.04
R8	100.00	91.67	83.33	100.00	75.00	100.00	100.00	100.00	75.00	91.67	91.74
R9	100.00	100.00	100.00	100.00	91.67	100.00	83.33	100.00	100.00	75.00	95.04
R10	83.33	100.00	91.67	83.33	100.00	91.67	100.00	84.62	100.00	100.00	93.39
Avr											93.72 ± 1.42
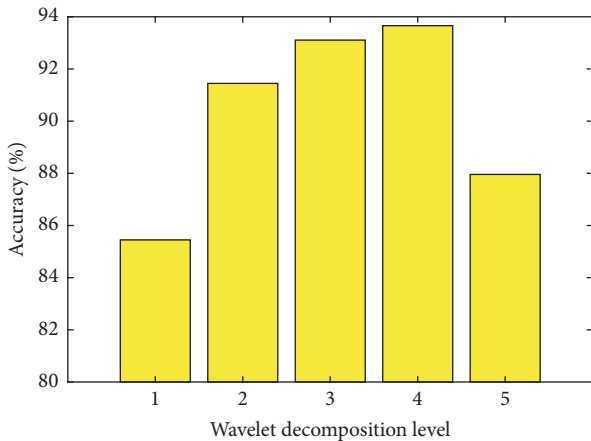
FIGURE 8: Choosing the optimal wavelet decomposition level.

corresponding accuracy varied as shown in Figure 8. Here, the accuracy is 85.45%, 91.45%, 93.11%, 93.66%, and 87.96% when decomposition level is 1, 2, 3, 4, and 5, respectively. Obviously, the 4th-level decomposition yields the greatest accuracy; hence, we chose the optimal decomposition level as 4. The Renyi entropy was then calculated over all the subbands of this 4-level wavelet decomposition.

Note that 3D-DWT is more straightforward than 2D-DWT over a one particular slice. Nevertheless, our aim in this study is to select a distinguishing slice, which is related to brain regions affected by alcoholism, in order to reduce the computation burden. In the future, we shall test the results of 3D-DWT.

6.4. Optimal Renyi Order. In this experiment, we shall illustrate why we set the Renyi order α as 1.2 by a coarse-to-fine grid search. First, we search the coarse grid from 0 to 6 with an increase of 1. The result was shown in Figure 9(a), and the value of 1 was selected as the initial point for fine grid search. Second, the fine grid from 0.5 to 1.5 with an increase of 0.1 was established, and the result was shown in Figure 9(b). We can observe that $\alpha = 1.2$ can yield the greatest accuracy.

6.5. Effectiveness of Three-Segment Encoding. Our proposed TSE-Jaya can train the weights, biases, and number of hidden neurons (NHN) simultaneously. In this experiment, we compare TSE-Jaya to plain Jaya algorithm [10], which can only train the weights and biases of feedforward neural network [10]; hence, we have to fix the number of hidden neurons by experience. Here, we set NHN as 10 for plain Jaya algorithm. The settings of other parameters were the same as previous experiments. The results of 10 × 10-fold cross validation of plain Jaya [10] are shown in Table 8, and the

Table 5: Accuracy of the proposed method (unit: %).

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Total
R1	100.00	95.65	100.00	91.30	100.00	91.67	91.30	95.83	95.83	91.30	95.32
R2	100.00	95.83	100.00	87.50	100.00	86.96	82.61	91.67	100.00	95.83	94.04
R3	100.00	100.00	95.65	100.00	100.00	100.00	95.83	100.00	78.26	79.17	94.89
R4	95.83	91.30	91.30	91.67	79.17	100.00	100.00	95.65	83.33	95.65	92.34
R5	95.65	82.61	100.00	95.83	95.83	95.83	82.61	86.96	95.65	91.67	92.34
R6	100.00	83.33	92.00	100.00	100.00	95.65	86.96	75.00	100.00	95.83	92.77
R7	95.65	91.67	86.96	95.65	86.96	91.30	96.00	100.00	100.00	95.65	94.04
R8	96.00	95.83	87.50	100.00	82.61	95.65	100.00	95.65	86.96	91.30	93.19
R9	95.65	100.00	100.00	100.00	95.65	95.65	86.96	95.83	100.00	83.33	95.32
R10	91.67	95.65	91.67	82.61	100.00	91.30	95.65	84.00	95.65	95.83	92.34
Avr											93.66 ± 1.23

Table 6: Results of 10 × 10-fold cross validation of four state-of-the-art methods.

FA + PNN [2]	Sen	Spc	Acc	IPSO [3]	Sen	Spc	Acc
R1	86.84	86.78	86.81	R1	90.35	87.60	88.94
R2	83.33	90.08	86.81	R2	92.11	85.95	88.94
R3	84.21	87.60	85.96	R3	89.47	90.08	89.79
R4	90.35	85.12	87.66	R4	88.60	90.08	89.36
R5	87.72	84.30	85.96	R5	88.60	88.43	88.51
R6	84.21	86.78	85.53	R6	88.60	90.08	89.36
R7	81.58	91.74	86.81	R7	89.47	90.91	90.21
R8	86.84	87.60	87.23	R8	85.96	93.39	89.79
R9	85.09	86.78	85.96	R9	87.72	90.91	89.36
R10	87.72	85.95	86.81	R10	87.72	86.78	87.23
Avr	85.79 ± 2.58	87.27 ± 2.21	86.55 ± 0.67	Avr	88.86 ± 1.66	89.42 ± 2.23	89.15 ± 0.83
HMI + SVM [5]	Sen	Spc	Acc	PZM [6]	Sen	Spc	Acc
R1	82.46	86.78	84.68	R1	87.72	86.78	87.23
R2	85.96	83.47	84.68	R2	87.72	87.60	87.66
R3	85.09	86.78	85.96	R3	85.09	90.91	88.09
R4	87.72	85.12	86.38	R4	85.96	86.78	86.38
R5	85.09	87.60	86.38	R5	85.96	85.95	85.96
R6	83.33	86.78	85.11	R6	89.47	87.60	88.51
R7	85.96	85.95	85.96	R7	85.09	83.47	84.26
R8	79.82	89.26	84.68	R8	85.96	88.43	87.23
R9	80.70	82.64	81.70	R9	85.96	85.95	85.96
R10	89.47	82.64	85.96	R10	83.33	86.78	85.11
Avr	84.56 ± 3.02	85.70 ± 2.21	85.15 ± 1.40	Avr	86.23 ± 1.71	87.02 ± 1.91	86.64 ± 1.35

comparison between Jaya [10] and our proposed TSE-Jaya is shown in Table 9.

The superiority of proposed TSE-Jaya to plain Jaya [10] is clear. This demonstrates the importance of choosing the optimal number of hidden neurons, that is, that the variable number of hidden neurons gives a better performance than fixed number of hidden neurons, which is also validated by Carleo and Troyer (2017) [35].

6.6. *Training Algorithm Comparison.* To demonstrate the efficiency of the proposed algorithm, we have compared the TSE-Jaya with a several global optimization algorithms including MOGA [11], PSO [12], BCO [13], MACS [14],

and RCBBO [15]. All the settings of common controlling parameters are the same: the maximum iterative epoch is set as 1000, and the population in all algorithms is set to 20. The algorithm-specific parameters of those four comparison algorithms are assigned by experiences. The results of those four training algorithms over 10 × 10-fold cross validation are shown in Table 10, and the final comparison with our proposed TSE-Jaya was shown in Table 11.

Table 11 shows that the proposed TSE-Jaya performed the best among all six algorithms. The PSO [12] and RCBBO [15] ranked the second and the third, with their accuracies over than 90%. The BCO [13] ranked the fourth, the MACS [14] ranked the fifth, and MOGA [11] performed the worst. The

TABLE 7: Method comparison.

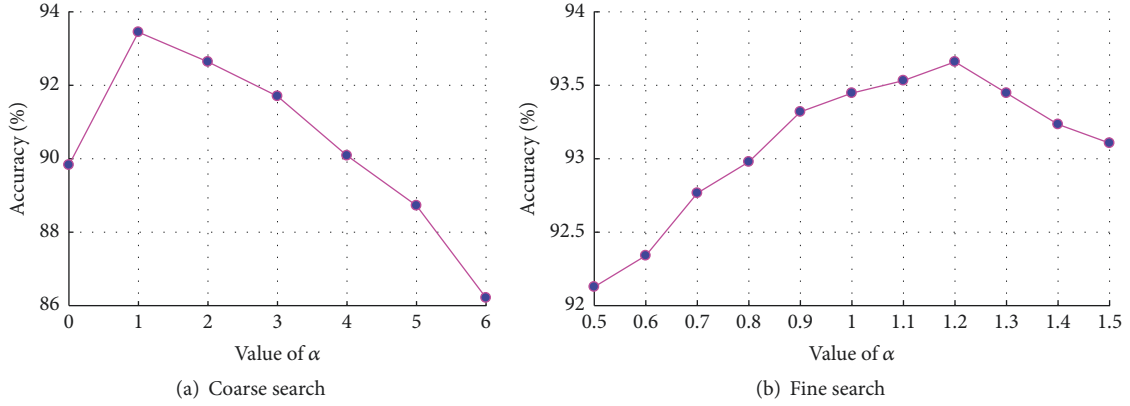| Alcoholism identification method	Sensitivity	Specificity	Accuracy
FA + PNN [2]	85.79 ± 2.58	87.27 ± 2.21	86.55 ± 0.67
IPSO [3]	88.86 ± 1.66	89.42 ± 2.23	89.15 ± 0.83
HMI + SVM [5]	84.56 ± 3.02	85.70 ± 2.21	85.15 ± 1.40
PZM [6]	86.23 ± 1.71	87.02 ± 1.91	86.64 ± 1.35
WMI + FNN + TSE-Jaya (ours)	93.60 ± 1.55	93.72 ± 1.42	93.66 ± 1.23

(a) Coarse search

(b) Fine search

FIGURE 9: Choosing the optimal Renyi order by coarse-to-fine grid search.

TABLE 8: Results of 10 × 10-fold cross validation of plain Jaya algorithm [10].

Jaya [10]	Sen (%)	Spc (%)	Acc (%)
R1	92.11	92.56	92.34
R2	94.74	95.04	94.89
R3	87.72	93.39	90.64
R4	92.11	91.74	91.91
R5	94.74	91.74	93.19
R6	96.49	94.21	95.32
R7	89.47	95.04	92.34
R8	92.11	92.56	92.34
R9	92.11	90.91	91.49
R10	95.61	92.56	94.04

TABLE 9: Plain Jaya algorithm versus the proposed TSE-Jaya.

Method	Sensitivity	Specificity	Accuracy
Jaya [10]	92.72 ± 2.74	92.98 ± 1.42	92.85 ± 1.50
TSE-Jaya (ours)	93.60 ± 1.55	93.72 ± 1.42	93.66 ± 1.23

reasons behind efficiency of the proposed approach can be explained from two aspects: (i) The Jaya does not need to set the algorithm-specific parameters, making it more reliable than other algorithms. (ii) The TSE guarantees the variable number of hidden neurons at each run.

6.7. Validation of the Selected Slice. In this experiment, we validated in terms of the classification performance the

selection of the 80th slice. We set a range of Z increasing from 30 to 150 with an increment of 10 as shown in Figure 10.

Other settings were the same as the previous experiments. Again, 10 repetitions of 10-fold cross validation were utilized. The curve of accuracy is drawn in Figure 11. It is observed that the 80th slice gives the highest accuracy among all candidate slices. The reason is that this slice contains the enlarged ventricle and the shrunk gray matter caused by alcoholism. On the contrary, hippocampus [36] and striatum [37] are also related to alcoholism. Nevertheless, their altered volumes are relatively small and hence do not provide an excellent performance in this task.

In this case, the optimal slice could be in a position that is vertical to x or y-axes, or it can be even an oblique plane to all three axes. Here, we choose a slice vertical to z-axis, which is for the convenience of radiologists, since they usually read the axial slices. In the future, we shall develop techniques to handle multislices, and we may develop surface analysis techniques [38].

7. Conclusions

In this study, we proposed a novel alcoholism identification method from healthy controls based on a computer-vision approach. Our method was based on three components: the proposed wavelet Renyi entropy, feedforward neural network, and the proposed three-segment encoded Jaya algorithm. The experiments showed that our method achieved a sensitivity of 93.60 ± 1.55%, a specificity of 93.72 ± 1.42, and an accuracy of 93.66 ± 1.23 over a 10 × 10-fold cross validation. The performance is superior to four state-of-the-art alcoholism algorithms. We validated the optimal wavelet

TABLE 10: Results of 10 × 10-fold cross validation of four training algorithms.

MOGA [11]	Sen	Spc	Acc	PSO [12]	Sen	Spc	Acc
R1	78.07	75.21	76.60	R1	90.35	90.91	90.64
R2	76.32	79.34	77.87	R2	87.72	87.60	87.66
R3	72.81	77.69	75.32	R3	87.72	90.08	88.94
R4	75.44	77.69	76.60	R4	92.98	93.39	93.19
R5	76.32	76.86	76.60	R5	88.60	90.08	89.36
R6	75.44	75.21	75.32	R6	89.47	87.60	88.51
R7	78.95	80.99	80.00	R7	88.60	87.60	88.09
R8	76.32	73.55	74.89	R8	93.86	90.91	92.34
R9	73.68	81.82	77.87	R9	90.35	92.56	91.49
R10	79.82	80.17	80.00	R10	92.98	94.21	93.62
Avr	76.32 ± 2.19	77.85 ± 2.72	77.11 ± 1.83	Avr	90.26 ± 2.28	90.50 ± 2.41	90.38 ± 2.18
BCO [13]	Sen	Spc	Acc	MACS [14]	Sen	Spc	Acc
R1	87.72	85.12	86.38	R1	84.21	85.95	85.11
R2	92.11	90.91	91.49	R2	86.84	89.26	88.09
R3	87.72	88.43	88.09	R3	81.58	85.95	83.83
R4	85.96	90.91	88.51	R4	84.21	79.34	81.70
R5	84.21	86.78	85.53	R5	80.70	83.47	82.13
R6	87.72	87.60	87.66	R6	82.46	85.12	83.83
R7	91.23	90.08	90.64	R7	83.33	84.30	83.83
R8	91.23	85.12	88.09	R8	84.21	86.78	85.53
R9	84.21	85.12	84.68	R9	86.84	86.78	86.81
R10	84.21	91.74	88.09	R10	88.60	86.78	87.66
Avr	87.63 ± 3.05	88.18 ± 2.61	87.91 ± 2.09	Avr	84.30 ± 2.50	85.37 ± 2.64	84.85 ± 2.19
RCBBO [15]	Sen	Spc	Acc				
R1	84.21	90.08	87.23				
R2	88.60	90.91	89.79				
R3	88.60	90.08	89.36				
R4	93.86	93.39	93.62				
R5	91.23	90.91	91.06				
R6	92.11	93.39	92.77				
R7	90.35	88.43	89.36				
R8	91.23	90.08	90.64				
R9	90.35	89.26	89.79				
R10	88.60	87.60	88.09				
Avr	89.91 ± 2.62	90.41 ± 1.88	90.17 ± 1.95				

TABLE 11: FNN training algorithm comparison.

FNN training algorithm	Sensitivity	Specificity	Accuracy	Rank
MOGA [11]	76.32 ± 2.19	77.85 ± 2.72	77.11 ± 1.83	6
PSO [12]	90.26 ± 2.28	90.50 ± 2.41	90.38 ± 2.18	2
BCO [13]	87.63 ± 3.05	88.18 ± 2.61	87.91 ± 2.09	4
MACS [14]	84.30 ± 2.50	85.37 ± 2.64	84.85 ± 2.19	5
RCBBO [15]	89.91 ± 2.62	90.41 ± 1.88	90.17 ± 1.95	3
TSE-Jaya (ours)	93.60 ± 1.55	93.72 ± 1.42	93.66 ± 1.23	1

decomposition to be 4, and the optimal Renyi order was 1.2. Besides, comparing to the existing global optimization approaches, the proposed three-segment encoded Jaya is proven to provide a better performance than other methods such as plain Jaya and another five training algorithms.

Finally, we validated the reason why we chose the 80th slice.

The shortcomings of our method lie in two aspects. First, our method needs to scan the whole brain and select the 80th slice at z-axis. Second, the wavelet Renyi entropy was
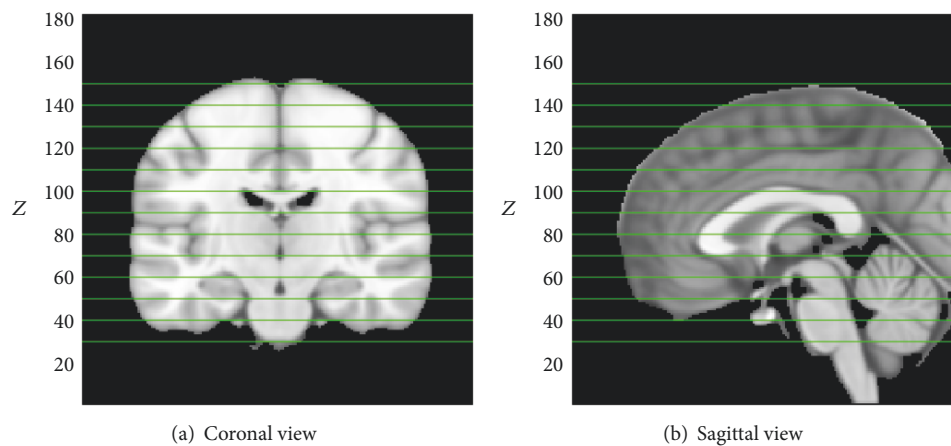
(a) Coronal view

(b) Sagittal view
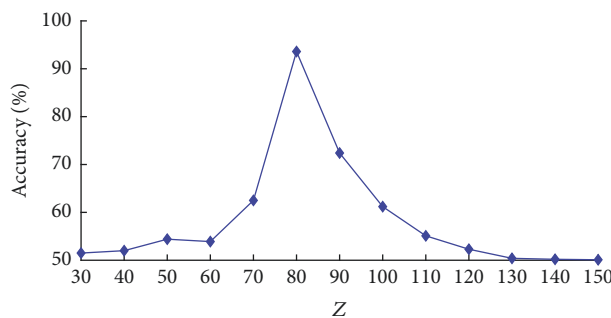
FIGURE 10: Position of Z-slice.

FIGURE 11: Classification performance changes with slice index Z.

extracted, but in the future we shall try to find more efficient features.

In the future work, we will develop classifiers based on multisource data, such as facial image, EEG, and spectrum data. We shall explore the changes of functional connectivity of alcoholism brain. Second, other image features, such as histogram of oriented gradient, will be tested. Third, deep learning methods will be tested.

Conflicts of Interest

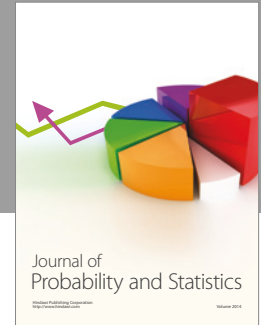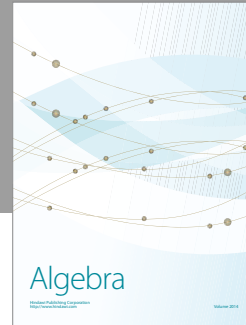The authors have no conflicts of interest to disclose with regard to the subject matter of this paper.
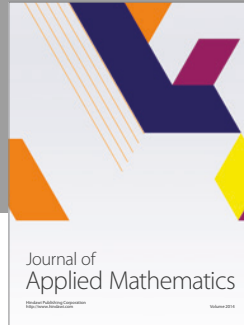
Acknowledgments

References

[1] D. R. Nayak, R. Dash, and B. Majhi, "Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests," *Neurocomputing*, vol. 177, pp. 188–197, 2016.

[2] M. Alweshah and S. Abdullah, "Hybridizing firefly algorithms with a probabilistic neural network for solving classification problems," *Applied Soft Computing*, vol. 35, pp. 513–524, 2015.

[3] Y.-D. Lv and X.-X. Hou, "Alcoholism detection by medical robots based on Hu moment invariants and predator–prey adaptive-inertia chaotic particle swarm optimization," *Computers & Electrical Engineering*, vol. 63, pp. 126–138, 2017.

[4] M. A. Monnig, "Observed power and projected sample sizes to detect white matter atrophy in neuroimaging of alcohol use disorders," *Alcoholism-Clinical And Experimental Research*, vol. 36, article 272A, 2012.

[5] J. Yang, "Pathological brain detection in MRI scanning via Hu moment invariants and machine learning," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 29, no. 2, pp. 299–312, 2017.

[6] Y. Jiang and W. Zhu, "Exploring a smart pathological brain detection method on pseudo Zernike moment," *Multimedia Tools and Applications*, pp. 1–16, 2017.

[7] Y.-D. Lv and Y. Sui, "Alcoholism detection by data augmentation and convolutional neural network with stochastic pooling," *Journal of Medical Systems*, vol. 42, no. 1, article 2, 2017.

[8] R. M. Stephenson, G. R. Naik, and R. Chai, "A system for accelerometer-based gesture classification using artificial neural networks," in *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '17)*, pp. 4187–4190, Seogwipo, South Korea, July 2017.

[9] A. Barik, R. K. Rai, and A. Chowdhury, "Alcohol use-related problems among a rural Indian population of West Bengal: an application of the Alcohol Use Disorders Identification Test (AUDIT)," *Alcohol and Alcoholism*, vol. 51, no. 2, pp. 215–223, 2016.

[10] P. Phillips, "Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm," *Neurocomputing*, vol. 272, pp. 668–676, 2017.

[11] E. Hajimani, M. Ruano, and A. Ruano, "An intelligent support system for automatic detection of cerebral vascular accidents

from brain CT images," *Computer Methods and Programs in Biomedicine*, vol. 146, pp. 109–123, 2017.

[12] R. M. Chen, S. C. Yang, and C. M. Wang, "MRI brain tissue classification using unsupervised optimized extenics-based methods," *Computers & Electrical Engineering*, vol. 58, pp. 489–501, 2017.

[13] S. Subramaniam and M. Radhakrishnan, "Neural network with bee colony optimization for MRI brain cancer image classification," *International Arab Journal of Information Technolog*, vol. 13, no. 1, pp. 118–124, 2016.

[14] G. S. Raghtate and S. S. Salankar, "Automatic brain MRI classification using modified ant colony system and neural network classifier," in *Proceedings of the 7th International Conference on Computational Intelligence and Communication Networks (CICN '15)*, pp. 1241–1246, Jabalpur, India, December 2015.

[15] P. Chen and S. Du, "Pathological brain detection via wavelet packet tsallis entropy and real-coded biogeography-based optimization," *Fundamenta Informaticae*, vol. 151, no. 1-4, pp. 275–291, 2017.

[16] M. W. Woolrich, S. Jbabdi, B. Patenaude et al., "Bayesian analysis of neuroimaging data in FSL," *NeuroImage*, vol. 45, no. 1, pp. S173–S186, 2009.

[17] S. M. Smith, M. Jenkinson, M. W. Woolrich et al., "Advances in functional and structural MR image analysis and implementation as FSL," *NeuroImage*, vol. 23, supplement 1, pp. S208–S219, 2004.

[18] T. Hatchard, O. Mioduszewski, C. Fall, A. Byron-Alhassan, P. Fried, and A. M. Smith, "Neural impact of low-level alcohol use on response inhibition: an fMRI investigation in young adults," *Behavioural Brain Research*, vol. 329, pp. 12–19, 2017.

[19] T. Schmidt, P. Roser, O. Ze, G. Juckel, B. Suchan, and P. Thoma, "Cortical thickness and trait empathy in patients and people at high risk for alcohol use disorders," *Psychopharmacology*, vol. 234, no. 23-24, pp. 3521–3533, 2017.

[20] P. Bach, E. Zois, S. Vollstädt-Klein et al., "Association of the alcohol dehydrogenase gene polymorphism rs1789891 with gray matter brain volume, alcohol consumption, alcohol craving and relapse risk," *Addiction Biology*, 2017.

[21] T. Koza and N. Karaboga, "Quadrature mirror filter bank design for mitral valve doppler signal using artificial bee colony algorithm," *Elektronika ir Elektrotechnika*, vol. 23, no. 1, pp. 57–62, 2017.

[22] G. Savić, M. Prokin, V. Rajović, and D. Prokin, "High-performance 1-D and 2-D inverse DWT 5/3 filter architectures for efficient hardware implementation," *Circuits, Systems and Signal Processing*, vol. 36, no. 9, pp. 3674–3701, 2017.

[23] N. Masi, "Rényi entropy for particle systems as an instrument to enlarge the Boltzmannian concept of entropy: some holographic perspectives," *International Journal of Modern Physics D: Gravitation, Astrophysics, Cosmology*, vol. 26, no. 10, Article ID 1750105, 11 pages, 2017.

[24] C. Y. Kee, S. G. Ponnambalam, and C. K. Loo, "Binary and multi-class motor imagery using Renyi entropy for feature extraction," *Neural Computing & Applications*, vol. 28, no. 8, pp. 2051–2062, 2017.

[25] J. M. Gorriz and J. Ramírez, "Wavelet entropy and directed acyclic graph support vector machine for detection of patients with unilateral hearing loss," *Frontiers in Computational Neuroscience*, vol. 10, article 160, 2016.

[26] D. A. Winkler and T. C. Le, "Performance of deep and shallow neural networks, the universal approximation theorem activity

cliffs, and QSAR," *Molecular Informatics*, vol. 36, no. 11, Article ID 1781141, 2017.

[27] R. V. Rao, "Jaya: a simple and new optimization algorithm for solving constrained and unconstrained optimization problems," *International Journal of Industrial Engineering Computations*, vol. 7, no. 1, pp. 19–34, 2016.

[28] P. Ocłoń, P. Cisek, M. Rerak et al., "Thermal performance optimization of the underground power cable system by using a modified Jaya algorithm," *International Journal of Thermal Sciences*, vol. 123, pp. 162–180, 2018.

[29] K. Yu, J. Liang, B. Qu, X. Chen, and H. Wang, "Parameters identification of photovoltaic models using an improved JAYA optimization algorithm," *Energy Conversion and Management*, vol. 150, pp. 742–753, 2017.

[30] R. Rao and K. More, "Optimal design and analysis of mechanical draft cooling tower using improved Jaya algorithm," *International Journal of Refrigeration*, vol. 82, pp. 312–324, 2017.

[31] A. Kaur, S. Sharma, and A. Mishra, "Sensing period adaptation for multiobjective optimisation in cognitive radio using Jaya algorithm," *IEEE Electronics Letters*, vol. 53, no. 19, pp. 1335-1336, 2017.

[32] R. V. Rao and A. Saroj, "Constrained economic optimization of shell-and-tube heat exchangers using elitist-Jaya algorithm," *Energy*, vol. 128, pp. 785–800, 2017.

[33] K. Abhishek, V. R. Kumar, S. Datta, and S. S. Mahapatra, "Application of JAYA algorithm for the optimization of machining performance characteristics during the turning of CFRP (epoxy) composites: comparison with TLBO, GA, and ICA," *Engineering with Computers*, vol. 33, no. 3, pp. 457–475, 2017.

[34] H. M. Lu, "Facial emotion recognition based on biorthogonal wavelet entropy, fuzzy support vector machine, and stratified cross validation," *IEEE Access*, vol. 4, pp. 8375–8385, 2016.

[35] G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks," *Science*, vol. 355, no. 6325, pp. 602–606, 2017.

[36] T. Murano, H. Koshimizu, H. Hagihara, and T. Miyakawa, "Transcriptomic immaturity of the hippocampus and prefrontal cortex in patients with alcoholism," *Scientific Reports*, vol. 7, Article ID 44531, 2017.

[37] M. Spoelder, P. Hesseling, M. Styles et al., "Dopaminergic neurotransmission in ventral and dorsal striatum differentially modulates alcohol reinforcement," *European Journal of Neuroscience*, vol. 45, no. 1, pp. 147–158, 2017.

[38] B. M. Craver, M. M. Acharya, B. D. Allen et al., "3D surface analysis of hippocampal microvasculature in the irradiated brain," *Environmental and Molecular Mutagenesis*, vol. 57, no. 5, pp. 341–349, 2016.