Advances in Intelligent Systems Research, volume 148
International Conference on Information Technology and Management Engineering (ICITME 2018)

Alcoholism Detection via Wavelet Energy and Logistic Regression

Yuanyuan Tao[1] and Felix Macdonald[2,*]
[1]School of Computer Science and Technology, Nanjing Normal University, Nanjing, Jiangsu 210023, China
[2]Université de Sherbrooke, Sherbrooke, Québec, QC J1K 2R1, Canada
*Corresponding author

***Abstract*—In this study, we proposed an application of alcoholism detection via wavelet energy and logistic regression. We collected data sets of 70 volunteers who signed up through advertising, among which 35 were with alcoholism and the rest were healthy. We first used wavelet energy (WN) to extract brain images features. Then, we employed logistic regression (LR) as the classification tool. Finally, we used 5-fold stratified cross validation to verify classifier performance. Our method achieves a sensitivity of 84.00± 3.86%, a specificity of 84.86± 3.03%, and an accuracy of 84.43± 1.42%. Our method gives better performance than HWT and ANN-GA.**

Keywords—alcoholism; wavelet energy; logistic regression; detection; identification

I. Introduction

Alcoholism, commonly known as drunkenness, is the effect of excessive consumption of alcohol on the central nervous system [1]. Alcohol mainly damages the central nervous system of the human body, which makes the function of the nervous system disorganized and suppressed, and the serious intoxication can lead to the death of the central inhibition and paralysis of the respiratory circulation center [2, 3]. Considering the disadvantages of traditional detection, we need a more advanced method to detect. Nowadays, image processing methods [4-6] are commonly and widely used in disease identification.

In the advance of our research, we found that there are many papers have discussed the algorithms of alcoholism detection. Lv (2018) [7] used convolutional neural network (CNN) with stochastic pooling. Hou (2017) [8] employed Hu moment invariants and predator-prey adaptive-inertia chaotic particle swarm optimization. Ahmadi, Pei (2017) [9] proposed EEG signals and functional brain network features extraction. Lima (2018) [10] used Haar wavelet transform (HWT) to identify alcoholism. Dil, Ghaedi (2016) [11] presented an artificial neural network genetic algorithm (ANN-GA).

In our study, we collected data from brain magnetic resonance imaging [12-18]. We proposed the method of Wavelet Energy (WN) to extract the feature from brain images. It can extract the brain image features and make the brain image classification more accurate. The wavelet energy of each wavelet sub-band is used to measure the discriminating potential brain image from the obtained texture features by implementing wavelet packet-based texture classification.

Then, logistic regression (LR) [19] was used to identify the brain image of alcoholism. Linear regression and logistic regression are relatively basic and frequently used contents in machine learning. Linear regression is mainly used to solve the problem of continuous value prediction. Logistic regression is used to solve the problem of classification, and the probability of output belongs to a certain category.

II. Materials and Methods

A. Materials

In total 70 volunteers were enrolled by advertisements. 35 are with alcoholism and the rest are healthy. Magnetic resonance imaging (MRI) [20-23] was performed over each subject. The method we proposed in this paper includes the following steps as shown in Figure 1.

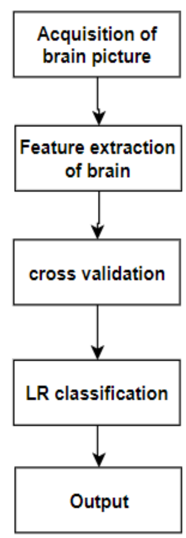


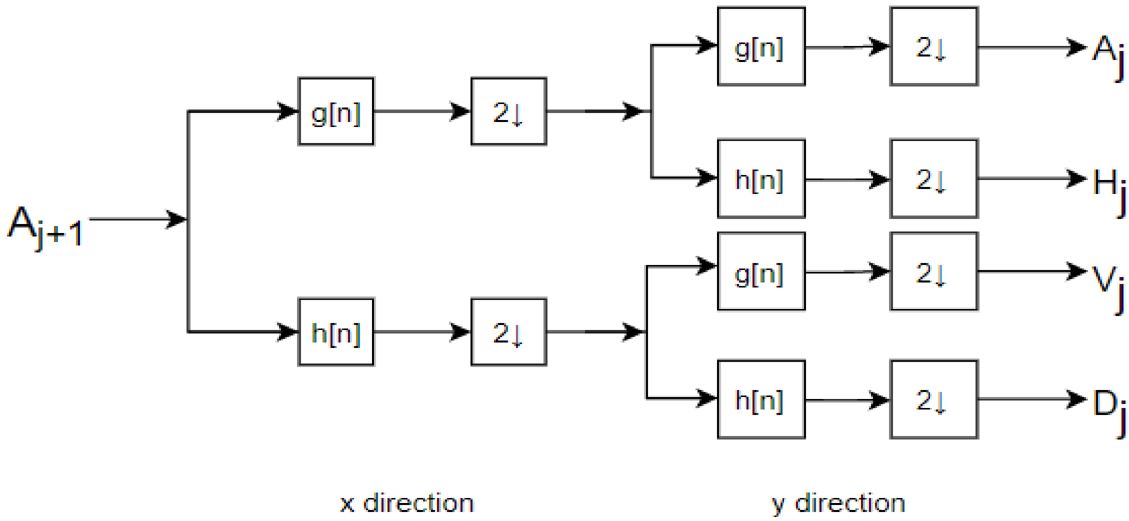
FIGURE I. DIAGRAM OF OUR METHOD

Alcoholism is a very troublesome disease. Its symptoms are mainly nausea and vomiting, dizziness, and headaches. Severe patients can cause respiratory depression, liver damage, and gastric ulcers.

B. Feature Extraction

If the original data is regarded as a two-dimensional discrete data after an image sampling, then the decomposition of the two-dimensional discrete wavelet transform (2D-DWT) [24] can be regarded as a filtering process as shown in Figure 2, that is, filtering the horizontal direction first, then filtering the vertical direction, and obtaining four different frequency bands

A_j,H_j,V_j and D_j,where A_j is the low frequency and H_j, V_j, D_j are the high frequency components on the horizontal, vertical, and diagonal.



(g[n] represents low pass filter and h[n] represents high pass filter)

FIGURE II. THE 2D DISCRETE WAVELET TRANSFORM

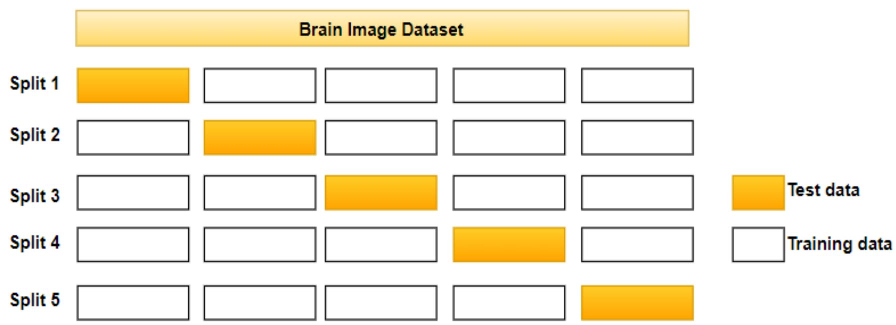


FIGURE III. THE FIVE-FOLD STRATIFIED CROSS VALIDATION

Many features can be extracted from the wavelet coefficients [25-27]. A simple and effective method for feature extraction from wavelet coefficients is the wavelet energy (WN) [28] of detail images. The WE in horizontal, vertical and diagonal directions of high frequency at *i*-th level can be, respectively, defined as:

$$E_i(H_j) = \sum_x \sum_y (H_j(x,y))^2 \quad (1)$$

$$E_i(V_j) = \sum_x \sum_y (V_j(x,y))^2 \quad (2)$$

$$E_i(V_j) = \sum_x \sum_y (V_j(x,y))^2 \quad (3)$$

where E represents the wavelet-energy. These energies reflect the strength of the images' details in different directions.

C. Logistic Regression

Logistic regression is a classical method in statistics, which is very closely related to classification**.** It is commonly used in the last layer of convolutional neural network [29-32]. Hence, we chose logistic regression as a classifier. As we all know, the formula of linear regression [33] is as follows:

$$z = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n = \theta^T x \quad (4)$$

where x is the input sample and θ is a parameter. For logical regression, its idea is also based on linear regression, and its formula is as follows:

$$h_\theta(x) = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-\theta^T X}} \quad (5)$$

where $h_\theta(X)$ is the value predicted by the parameters θ and x.

D. Stratified Cross Validation

We used a 5-fold stratified cross validation, and each fold contains 7 alcoholism brains and 7 healthy control brains. The usage of stratified 5-fold is similar to 5-fold, but it is stratified sampling to ensure that the proportion of different types of samples in the training set and test set is the same as that in the original data set [34-36]. As in Figure 3, each fold has the same class distribution as the original dataset.

III. Experiment and Results

A. Cross Validation Result

The experiment was performed by in-house software, and run on the platform of Matlab 2014a. All the parameters were tuned and secured by trial-and-error method. We report the sensitivity, specificity, and accuracy of all runs in Table 1. As is shown, the sensitivity of our method is 84.00± 3.86, the specificity is 84.86± 3.03, and the accuracy is 84.43± 1.42.

TABLE I. PERFORMANCE OF PROPOSED METHOD

Run	Sensitivity	Specificity	Accuracy
1	80.00	88.57	84.29
2	85.71	88.57	87.14
3	77.14	88.57	82.86
4	85.71	85.71	85.71
5	82.86	82.86	82.86
6	85.71	82.86	84.29
7	91.43	80.00	85.71
8	82.86	85.71	84.29
9	85.71	82.86	84.29
10	82.86	82.86	82.86
Average	84.00± 3.86	84.86± 3.03	84.43± 1.42

B. Algorithm Comparison

We compared our method with HWT [10] and ANN-GA [11]. The performances were performed over our dataset, and the results are presented in Table 2. Figure 4 shows the comparison plot. Our WN-LR method yields the greatest sensitivity, specificity and accuracy.

TABLE II. ALGORITHM COMPARISON

Approach	Sensitivity	Specificity	Accuracy
HWT [10]	81.71±4.51	81.43±4.52	81.57±2.18
ANN-GA [11]	76.00±3.07	77.14±5.22	76.57±1.54
WN-LR (Our)	84.00± 3.86	84.86±3.03	84.43±1.42

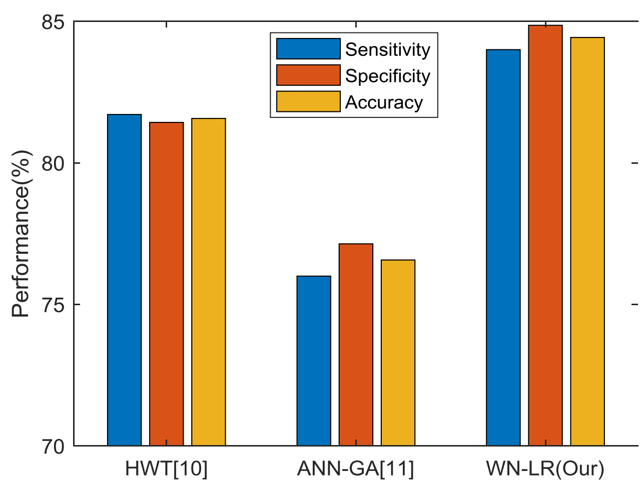


FIGURE IV. ALGORITHM COMPARISON

C. Decomposition Level Comparison

Our previous method used a decomposition level (L) of 2. Now we compared using decomposition levels of 1, 3, and 4. The results are listed in Table 3. We can find the performance achieved the best at the condition where L = 2. Figure 5 shows there is a significant increase from L = 1 to L = 2, and then gradually decrease from L = 2 to L = 4.

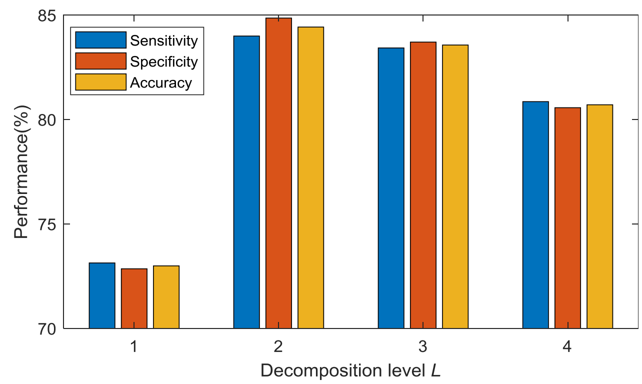


FIGURE V. DECOMPOSITION LEVEL COMPARISON

IV. Conclusion

In this paper, we proposed a new approach for automatic alcoholism detection of brain images using WE and LR. The results showed that the proposed method gave better performance than HWT and ANN-GA. However, our method also has some shortcomings as follows: (i) our dataset is small. (ii) We did not use the most popular deep learning method.

In the future, we may focus on following regards: (i)we shall either collect more brain images or use data augmentation techniques. (ii) Some advanced pattern recognition techniques may be used, such as convolutional neural network(CNN).

TABLE III. COMPARISON OF DECOMPOSITION LEVELS

Run	$L = 1$			$L = 2$			$L = 3$			$L = 4$		
	Sen	Spc	Acc	Sen	Spc	Acc	Sen	Spc	Acc	Sen	Spc	Acc
1	77.14	71.43	74.29	80.00	88.57	84.29	85.71	82.86	84.29	82.86	80.00	81.43
2	74.29	71.43	72.86	85.71	88.57	87.14	85.71	80.00	82.86	77.14	85.71	81.43
3	68.57	80.00	74.29	77.14	88.57	82.86	85.71	77.14	81.43	80.00	82.86	81.43
4	65.71	80.00	72.86	85.71	85.71	85.71	80.00	91.43	85.71	80.00	77.14	78.57
5	74.29	68.57	71.43	82.86	82.86	82.86	85.71	80.00	82.86	77.14	88.57	82.86
6	74.29	74.29	74.29	85.71	82.86	84.29	85.71	80.00	82.86	85.71	77.14	81.43
7	71.43	74.29	72.86	91.43	80.00	85.71	82.86	85.71	84.29	82.86	80.00	81.43
8	74.29	71.43	72.86	82.86	85.71	84.29	85.71	82.86	84.29	77.14	85.71	81.43
9	77.14	68.57	72.86	85.71	82.86	84.29	77.14	91.43	84.29	80.00	77.14	78.57
10	74.29	68.57	71.43	82.86	82.86	82.86	80.00	85.71	82.86	85.71	71.43	78.57
Average	73.14 ± 3.61	72.86 ± 4.31	73.00 ± 1.05	84.00 ± 3.86	84.86 ± 3.03	84.43 ± 1.42	83.43 ± 3.24	83.71 ± 4.87	83.57 ± 1.21	80.86 ± 3.31	80.57 ± 5.18	80.71 ± 1.54

REFERENCES

[1]. Stojakovic, A., et al., Several behavioral traits relevant for alcoholism are controlled by gamma 2 subunit containing GABA(A) receptors on dopamine neurons in mice. Neuropsychopharmacology, 2018. 43(7): p. 1548-1556

[2]. Han, L., *Identification of Alcoholism based on wavelet Renyi entropy and three-segment encoded Jaya algorithm.* Complexity, 2018. **2018**: Article ID. 3198184

[3]. Qian, P., *Cat Swarm Optimization applied to alcohol use disorder identification*. Multimedia Tools and Applications, 2018, doi: 10.1007/s11042-018-6003-8.

[4]. Wu, L.N., *Segment-based coding of color images.* Science in China Series F-Information Sciences, 2009. **52**(6): p. 914-925

[5]. Wei, G., *Color Image Enhancement based on HVS and PCNN.* SCIENCE CHINA Information Sciences, 2010. **53**(10): p. 1963-1976

[6]. Wu, L.N., *Improved image filter based on SPCNN.* Science In China Series F-Information Sciences, 2008. **51**(12): p. 2115-2125

[7]. Lv, Y.-D., *Alcoholism detection by data augmentation and convolutional neural network with stochastic pooling.* Journal of Medical Systems, 2018. **42**(1): Article ID. 2

[8]. Hou, X.-X., *Alcoholism detection by medical robots based on Hu moment invariants and predator-prey adaptive-inertia chaotic particle swarm optimization.* Computers and Electrical Engineering, 2017. **63**: p. 126-138

[9]. Ahmadi, N., et al. *Detection of Alcoholism Based on EEG Signals and Functional Brain Network Features Extraction*. in *30th International Symposium on Computer-Based Medical Systems (CBMS)*. 2017. Thessaloniki, Greece: IEEE. p. 179-184

[10]. Lima, D., Alcoholism detection in magnetic resonance imaging by Haar wavelet transform and back propagation neural network. AIP Conference Proceedings, 2018. 1955(1): Article ID. 040012

[11]. Dil, E.A., et al., Trace determination of safranin O dye using ultrasound assisted dispersive solid-phase micro extraction: Artificial neural network-genetic algorithm and response surface methodology. Ultrasonics Sonochemistry, 2016. 33: p. 129-140

[12]. Yang, J., *Pathological brain detection in MRI scanning via Hu moment invariants and machine learning.* Journal of Experimental & Theoretical Artificial Intelligence, 2017. **29**(2): p. 299-312

[13]. Jiang, Y., *Exploring a smart pathological brain detection method on pseudo Zernike moment*. Multimedia Tools and Applications, 2017, doi: 10.1007/s11042-017-4703-0.

[14]. Chen, S., et al., *Magnetic resonance brain image classification based on weighted-type fractional Fourier transform and nonparallel support vector machine.* International Journal of Imaging Systems and Technology, 2015. **25**(4): p. 317-327

[15]. Wang, S., *A novel method for magnetic resonance brain image classification based on adaptive chaotic PSO.* Progress in Electromagnetics Research, 2010. **109**: p. 325-343

[16]. Wu, L., *Magnetic Resonance Brain Image Classification by an Improved Artificial Bee Colony Algorithm.* Progress in Electromagnetics Research, 2011. **116**: p. 65-79

[17]. Peterson, B.S., *A Two-Level Iterative Reconstruction Method for Compressed Sensing MRI.* Journal of Electromagnetic Waves and Applications, 2011. **25**(8-9): p. 1081-1091

[18]. Wu, L., *An MR brain images classifier via principal component analysis and kernel support vector machine.* Progress In Electromagnetics Research, 2012. **130**: p. 369-388

[19]. Sadatrasoul, S.M., *Matrix Sequential Hybrid Credit Scorecard Based on Logistic Regression and Clustering.* Iranian Journal of Management Studies, 2018. **11**(1): p. 91-111

[20]. Bradley, P.S., *A support-based reconstruction for SENSE MRI.* Sensors, 2013. **13**(4): p. 4029-40

[21]. Ji, G., *An MR brain images classifier system via particle swarm optimization and kernel support vector machine.* The Scientific World Journal, 2013: Article ID. 130134

[22]. Wang, S., et al., *An improved quality guided phase unwrapping method and its applications to MRI.* Progress In Electromagnetics Research, 2014. **145**: p. 273-286

[23]. Dong, Z., *Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree.* Progress In Electromagnetics Research, 2014. **144**: p. 171-184

[24]. Bhattacharyya, A., et al., *Fourier-Bessel series expansion based empirical wavelet transform for analysis of non-stationary signals.* Digital Signal Processing, 2018. **78**: p. 185-196

[25]. Li, Y.-J., *Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization.* Multimedia Tools and Applications, 2018. **77**(9): p. 10393-10417

[26]. Phillips, P., *Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm.* Neurocomputing, 2018. **272**: p. 668-676

[27]. Li, P., et al., *Pathological Brain Detection via Wavelet Packet Tsallis Entropy and Real-Coded Biogeography-based Optimization.* Fundamenta Informaticae, 2017. **151**(1-4): p. 275-291

[28]. Goksu, H., *Ground moving target recognition using log energy entropy of wavelet packets.* Electronics Letters, 2018. **54**(4): p. 233-U48

[29]. Jia, W.-J. *Ford Motorcar Identification from Single-Camera Side-View Image Based on Convolutional Neural Network*. in *18th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL)*. 2017. Guilin, China: Springer. p. 173-180

[30]. Pan, C., *Abnormal breast identification by nine-layer convolutional neural network with parametric rectified linear unit and rank-based stochastic pooling.* Journal of Computational Science, 2018. **27**: p. 57-68

[31]. Muhammad, K., *Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation*. Multimedia Tools and Applications, 2017, doi: 10.1007/s11042-017-5243-3.

[32]. Zhao, G., *Polarimetric synthetic aperture radar image segmentation by convolutional neural network using graphical processing units*. Journal of Real-Time Image Processing, 2017, doi: 10.1007/s11554-017-0717-0.

[33]. Staatz, C., et al., *Evaluation of Multiple Linear Regression Limited Sampling Strategies for Enteric-Coated Mycophenolate Sodium in Adult Kidney Transplant Recipients.* American Journal of Transplantation, 2018. **18**: p. 802-802

[34]. Lu, H.M., *Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy, Fuzzy Support Vector Machine, and Stratified Cross Validation.* IEEE Access, 2016. **4**: p. 8375-8385

[35]. Nayak, D.R., *Detection of unilateral hearing loss by Stationary Wavelet Entropy.* CNS & Neurological Disorders - Drug Targets, 2017. **16**(2): p. 15-24

[36]. Yang, J.F., et al., *Magnetic resonance brain classification by a novel binary particle swarm optimization with mutation and time-varying acceleration coefficients.* Biomedical Engineering-Biomedizinische Technik, 2016. **61**(4): p. 431-441