Hearing Loss Detection in Medical Multimedia Data by Discrete Wavelet Packet Entropy and Single-Hidden Layer Neural Network Trained by Adaptive Learning-Rate Back Propagation

Shuihua Wang¹, Sidan Du², Yang Li², Huimin Lu³, Ming Yang⁴, Bin Liu⁵, and Yudong Zhang^{1,6(⊠)}

 ¹ School of Computer Science and Technology, Nanjing Normal University, Nanjing 210023, Jiangsu, China zhangyudong@njnu.edu.en
² School of Electronic Science and Engineering, Nanjing University, Nanjing 210046, Jiangsu, China
³ Department of Mechanical and Control Engineering, Kyushu Institute of Technology, Fukuoka Prefecture 804-8550, Japan
⁴ Department of Radiology, Children's Hospital of Nanjing Medical University, Nanjing 210008, China
⁵ Department of Radiology, Zhong-Da Hospital of Southeast University, Nanjing 210009, China
⁶ Translational Imaging Division, Columbia University, New York, NY 10032, USA

Abstract. In order to develop an efficient computer-aided diagnosis system for detecting left-sided and right-sided sensorineural hearing loss, we used artificial intelligence in this study. First, 49 subjects were enrolled by magnetic resonance imaging scans. Second, the discrete wavelet packet entropy (DWPE) was utilized to extract global texture features from brain images. Third, single-hidden layer neural network (SLNN) was used as the classifier with training algorithm of adaptive learning-rate back propagation (ALBP). The 10 times of 5-fold cross validation demonstrated our proposed method yielded an overall accuracy of 95.31%, higher than standard back propagation method with accuracy of 87.14%. Besides, our method also outperforms the "FRFT + PCA (Yang, 2016)", "WE + DT (Kale, 2013)", and "WE + MRF (Vasta 2016)". In closing, our method is efficient.

Keywords: Hearing loss · Multimedia data · Discrete wavelet packet entropy · Single-hidden layer neural network

1 Introduction

Multimedia data is a combination content of different data forms: text, audio, image, animation, and video. In medical application, the multimedia data offer refers to 3D volumetric data obtained by different imaging techniques.

The sensorineural hearing loss (SNHL) is a disease featuring in gradual deafness [1]. SNHL contains thee types: (i) sensory hearing loss (SHL), (ii) neural hearing loss (NHL), and (iii) both. SHL may be due to bad function of cochlear hair cell, and NHL may be because of impairment of cochlear nerve function.

In this study, we aimed to use multimedia data obtained by magnetic resonance imaging (MRI) scanning [2] to differentiate left-sided SNHL and right-sided SNHL. The detection basis is that SNHL patients will have slight to severe structural change in specific brain regions. Traditionally, the human eye-based detection is unreliable since the human eyes cannot perceive slight atrophy. Thus, artificial intelligence is employed in this study, which is aimed to develop a computer-aided diagnosis (CAD) system.

Traditional CAD systems mainly used discrete wavelet transform (DWT) [3–5] to learn global image features, and then employed latest pattern recognition tools. For example, Mao, Ma and Tian [6] used DWT to analyze the potential signals of local field. Ikawa [7] employed DWT to performance auditory brainstem response (ABR) operation. Nayak, Dash and Majhi [8] employed the DWT to identify brain images. They used AdaBoost with random forests as classifiers. Lahmiri [9] utilized three multi-resolution techniques: DWT, empirical mode decomposition (EMD), and variational mode decomposition (VMD). Chen and Chen [10] used principal component analysis (PCA) and generalized eigenvalue proximal support vector machine (GEPSVM). Gorriz and Ramírez [11] proposed a directed acyclic graph support vector machine method.

Nevertheless, DWT suffers from the disadvantage translational variance [12]. That means, even a slight translation may lead to different decomposition result [13]. Besides, the DWT decomposition will lead to larger dimension space ($\sim 10^6$) than original image ($\sim 10^5$) for a 256 × 256 size image, and it needs dimension reduction techniques, such as principal component analysis [14].

To solve this problem, we introduced a relatively new technique: discrete wavelet packet entropy (DWPE) [15–17] that can yield mere a few ($\sim 10^1$) translational invariant features. Besides, we used a single-hidden layer neural network as the classifier, which was trained by gradient descent with adaptive learning rate back propagation method.

2 Materials

Subjects were enrolled from outpatients of department of otorhinolaryngology and head-neck surgery and community. They were excluded if evidence existed of known psychiatric or neurological diseases, brain lesions, taking psychotropic medications, as well as contraindications to MR imaging.

Finally, the study collection includes 15 patients with left-sided SNHL (LSNHL), 14 patients with right-sided SNHL (RSNHL) and 20 age- and sex-matched healthy controls (HC), as shown in Table 1.

Preprocessing was implemented on the software platform of FMRIB Software Library (FSL) v5.0. The brain extraction tool (BET) was utilized to extract brain tissues. The results are shown in Fig. 1. Then, the extracted brains of all subjects were registered to MNI space. Three experienced radiologists were instructed to select the most distinctive (around 40-th) slice between SNHLs and HCs.

	HC	LSNHL	RSNHL
Gender (f/m)	12/8	7/8	8/6
Education level (year)	11.5 ± 3.2	12.5 ± 1.7	12.1 ± 2.4
Age (year)	53.6 ± 5.4	51.7 ± 9.6	53.9 ± 7.6
Disease duration (year)	-	17.6 ± 17.3	14.2 ± 14.9
PTA of left ear (dB)	22.2 ± 2.1	78.1 ± 17.9	21.8 ± 3.2
PTA of right ear (dB)	21.3 ± 2.2	20.4 ± 4.2	80.9 ± 17.4

Table 1. Subject characteristics

(PTA = pure tone average)



(c) sagittal direction

Fig. 1. The green lines label the edge of BET result (Color figure online)

3 Methodology

3.1 **Discrete Wavelet Packet Transform**

In the field of signal processing, standard discrete wavelet transform (abbreviated as DWT) [18, 19] decomposes the given signal at each level, by submitting the previous approximation subband to the quadrature mirror filters (QMF) [20]. Its even-indexed downsampling causes the translational invariance problem [21].

On the other hand, discrete wavelet packet transform (DWPT) [22] is an improvement of standard DWT. DWPT passes both approximation and detail coefficients of previous decomposition level to QMF, so it can create a full binary tree [23]. In general, DWPT offers more features than DWT at the same decomposition levels [24].

Suppose x represents the original signal, c the channel index, d the decomposition level, p the position parameter, D the decomposition coefficients, and ψ the wavelet function, then DWPT is calculated as below:

$$D_p^{c,d} = \int_{-\infty}^{\infty} x(t)\psi_c(2^{-d}t - p)\mathrm{d}t \tag{1}$$

where. 2^d sequences will be yielded. Based on *d*-level decomposition, the decomposition results of (d + 1) level is:

$$D_k^{2c,d+1} = \sum_{p \in \mathbb{Z}} h(p-2k) \times D_p^{c,d}, D_k^{2c+1,d+1} = \sum_{p \in \mathbb{Z}} l(p-2k) \times D_p^{c,d}$$
(2)



Fig. 2. Comparison between 2-level DWT and 2-level DWPT (x denotes for an image, H denotes the high-pass filter result, L denotes the low-pass filter result)

From Fig. 2, we can observe that for an image, DWT offer in total (1 + 3d) coefficient subbands. In contrast, DWPT generates in total 4^d coefficients subbands. Thus, DWPT can provide much more information than DWT.

3.2 Shannon Entropy

Entropy was originally utilized to measure the system disorder degree [25]. It was generalized by Shannon to measure information contained in a given message [26]. Suppose *m* the index of grey level, h_m the probability of *m*-th grey level, and *T* the total number of grey levels, we have the Shannon entropy *S* as:

$$S = -\sum_{m=1}^{T} h_m \log_2(h_m) \tag{3}$$

In the case of h_m equals to zero, the value of $0\log_2(0)$ is taken to 0 [27]. We calculated Shannon entropies of all subbands obtained from DWPT, and dubbed the results as discrete wavelet packet entropy (DWPE). For a brain image with size of 256 × 256, it has originally 65,536 features. A two-level DWPE can finally reduce the 65,536 features to only $2^4 = 16$ features.

3.3 Single-Hidden Layer Neural Network

The features were then presented into a classifier. There are many classifier in various fields, such as logistic regression [28], linear regression classifier [29, 30], extreme learning machine [31], decision tree [32, 33], etc.

In this study, we chose the classifier as a single-hidden layer neural-network (SLNN) [34] due to its superior performance. We did not employ multiple hidden layers [35], because one-hidden layer model is complicated enough to express our data. In a SLNN, the input nodes are connected to the hidden neuron layer, which is then connected to the output neuron layer.

The hidden neuron number is usually assigned with a large value. Afterwards, its value is decreased gradually till the classification performance reaches the peak result. The gradient descent with adaptive learning-rate back propagation (ALBP) algorithm [36] was employed to train the weights and biases of SLNN. Initial learning rate was set to 0.01. The increasing ratio and decreasing ratio of learning rate were set to 1.05 and 0.07, respectively. The maximum epoch is set to 5000.

4 **Experiments and Results**

(a) LSNHL Image (b) 1-level decomposition (c) 2-level decomposition

4.1 **DWPT Result**

Fig. 3. DWPT of a left-sided sensorineural hearing loss image

The 2-level DWPT result of a left-sided SNHL image is shown in Fig. 3. Here we can see in total 4 subbands are generated for 1-level decomposition, and 16 subbands are generated for 2-level decomposition.

4.2 **Accuracy Performance**

We repeated 5-fold cross validation [37] 10 times. The brief accuracy performance by BP algorithm is shown in left side in Table 2 with overall accuracy of 87.14%, and the accuracy performance by ALBP algorithm is shown in right side in Table 2 with overall accuracy of 95.31%. In these two tables, y/z represents y instances are successfully detected out of z instances.

The 10 repetition of 5-fold cross validation results indicate that this proposed ALBP performs better than classical BP algorithm. The reason lies in the adaptive learning-rate can accelerate the training procedure [38]. In standard BP, the learning



BP	F1	F2	F3	F4	F5	Т	Acc.	ALBP	F1	F2	F3	F4	F5	Т	Acc.
R1	10/10	9/10	7/9	10/10	8/10	44/49	89.80	R1	9/10	9/9	10/10	9/10	9/10	46/49	93.88
R2	10/10	10/10	9/10	7/9	8/10	44/49	89.80	R2	8/9	10/10	10/10	10/10	10/10	48/49	97.96
R3	9/9	9/10	9/10	8/10	10/10	45/49	91.84	R3	10/10	10/10	10/10	9/9	9/10	48/49	97.96
R4	6/10	10/10	7/9	9/10	7/10	39/49	79.59	R4	10/10	10/10	8/9	10/10	10/10	48/49	97.96
R5	9/10	9/10	8/9	8/10	8/10	42/49	85.71	R5	9/10	8/9	10/10	9/10	9/10	45/49	91.84
R6	9/10	9/9	8/10	7/10	9/10	42/49	85.71	R6	7/10	9/10	9/9	10/10	10/10	45/49	91.84
R7	8/10	8/10	9/10	9/9	10/10	44/49	89.80	R7	10/10	10/10	7/9	10/10	10/10	47/49	95.92
R8	10/10	10/10	9/10	7/9	9/10	45/49	91.84	R8	8/9	9/10	10/10	9/10	9/10	45/49	91.84
R9	7/10	9/10	9/10	9/9	6/10	40/49	81.63	R9	9/10	10/10	10/10	9/9	10/10	48/49	97.96
R10	9/9	8/10	9/10	9/10	7/10	42/49	85.71	R10	8/9	10/10	10/10	9/10	10/10	47/49	95.92
Т							87.14	Т							95.31

Table 2. Accuracy performance by BP and ALBP (R = Run; F = Fold; T = Total)

rate is unchanged, and thus the performance is sensitive to initial weight [39]. We see from left side of Table 2 that the accuracy in each run of BP vary from 79.59% to 91.84%. While the ALBP makes the learning rate responsive to the local error surface, and thus it is not as sensitive as BP. We see from right side of Table 2 that the accuracy in each run of ALBP vary from 91.84 to 97.76%. Thus, ALBP is much more stable than BP.

4.3 Comparison

Finally, we compared our DWPE + SLNN + ALBP approach with following three methods: (i) The combination of fractional Fourier transform (FRFT) and principal component analysis (PCA) method [40], which shall be abbreviated as FRFT + PCA. (ii) The combination of wavelet entropy (WE) and decision tree (DT) method [41], which is abbreviated as WE + DT. (iii) The hybrid system based on wavelet entropy (WE) and Markov random field (MRF) [42], abbreviated as WE + MRF.

Method	Overall accuracy
FRFT + PCA [40]	95.10%
WE + DT [41]	91.84%
WE + MRF [42]	91.02%
DWPE + SLNN + ALBP (Our)	95.31%

Table 3. Comparison with state-of-the-art methods

Table 3 shows that our method get superior overall accuracy of 95.31% to other three methods: FRFT + PCA [40], WE + DT [41], and WE + MRF [42]. The reason may be two folds: First, our method used DWPE, which combines two successful components, DWPT and Shannon entropy. Second, the wavelet packet transform is more efficient than fractional Fourier transform in image texture extraction. In the future, we shall try to use advanced classifiers, such as sparse autoencoder [43], convolutional neural network [44], and shared-weight neural network [45].

5 Conclusions

We developed a new computer-aided diagnosis system in this paper for detecting unilateral hearing loss, viz., left-sided or right-sided. The experiments gave promising results. In the future, we shall collect more data to further validate our method.

Acknowledgment. This study is supported by NSFC (61602250, 61271231), Program of Natural Science Research of Jiangsu Higher Education Institutions (16KJB520025, 15KJB470010), Natural Science Foundation of Jiangsu Province (BK20150983), Leading Initiative for Excellent Young Researcher (LEADER) of Ministry of Education, Culture, Sports, Science and Technology, Japan (16809746), Open Program of Jiangsu Key Laboratory of 3D Printing Equipment and Manufacturing (3DL201602), Open fund of Key Laboratory of Guangxi High Schools Complex System and Computational Intelligence (2016CSCI01), Open fund of Key Laboratory of Guangxi High Schools Complex System and Computational Intelligence (2016CSCI01).

References

- Nakagawa, T., Yamamoto, M., Kumakawa, K., Usami, S., Hato, N., Tabuchi, K., Takahashi, M., Fujiwara, K., Sasaki, A., Komune, S., Yamamoto, N., Hiraumi, H., Sakamoto, T., Shimizu, A., Ito, J.: Prognostic impact of salvage treatment on hearing recovery in patients with sudden sensorineural hearing loss refractory to systemic corticosteroids: a retrospective observational study. Auris Nasus Larynx 43, 489–494 (2016)
- Prasad, A., Ghosh, P.K.: Information theoretic optimal vocal tract region selection from real time magnetic resonance images for broad phonetic class recognition. Comput. Speech Lang. 39, 108–128 (2016)
- 3. Sun, P.: Preliminary research on abnormal brain detection by wavelet-energy and quantum-behaved PSO. Technol. Health Care **24**, S641–S649 (2016)
- Pattanaworapan, K., Chamnongthai, K., Guo, J.M.: Signer-independence finger alphabet recognition using discrete wavelet transform and area level run lengths. J. Vis. Commun. Image Represent. 38, 658–677 (2016)
- Dong, Z., Phillips, P., Ji, G., Yang, J.: Exponential wavelet iterative shrinkage thresholding algorithm for compressed sensing magnetic resonance imaging. Inf. Sci. 322, 115–132 (2015)
- Mao, Y., Ma, M.F., Tian, X.: Phase Synchronization Analysis of theta-band of Local Field Potentials in the Anterior Cingulated Cortex of Rats under Fear Conditioning. In: International Symposium on Intelligent Information Technology Application, pp. 737–741. IEEE Computer Society (2008)
- Ikawa, N.: Automated averaging of auditory evoked response waveforms using wavelet analysis, Int. J. Wavelets Multiresolut. Inf. Process. 11 (2013). Article ID: 1360009
- Nayak, D.R., Dash, R., Majhi, B.: Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests. Neurocomputing 177, 188– 197 (2016)
- 9. Lahmiri, S.: Image characterization by fractal descriptors in variational mode decomposition domain: application to brain magnetic resonance. Phys. A **456**, 235–243 (2016)

- Chen, Y., Chen, X.-Q.: Sensorineural hearing loss detection via discrete wavelet transform and principal component analysis combined with generalized eigenvalue proximal support vector machine and Tikhonov regularization. Multimedia Tools Appl. (2016). doi:10.1007/ s11042-016-4087-6
- Gorriz, J.M., Ramírez, J.: Wavelet entropy and directed acyclic graph support vector machine for detection of patients with unilateral hearing loss in MRI scanning, Frontiers Comput. Neurosci. 10 (2016) Article ID: 160
- Liu, A.: Magnetic resonance brain image classification via stationary wavelet transform and generalized eigenvalue proximal support vector machine. J. Med. Imaging Health Inform. 5, 1395–1403 (2015)
- Zhou, X.-X., Yang, J.-F., Sheng, H., Wei, L., Yan, J., Sun, P.: Combination of stationary wavelet transform and kernel support vector machines for pathological brain detection. Simulation 92, 827–837 (2016)
- 14. Ghods, A., Lee, H.H.: Probabilistic frequency-domain discrete wavelet transform for better detection of bearing faults in induction motors. Neurocomputing **188**, 206–216 (2016)
- Sun, Y.X., Zhuang, C.G., Xiong, Z.H.: Real-time chatter detection using the weighted wavelet packet entropy. In: IEEE/ASME International Conference on Advanced Intelligent Mechatronics, pp. 1652–1657. IEEE, New York (2014)
- Vyas, B., Maheshwari, R.P., Das, B.: Investigation for improved artificial intelligence techniques for thyristor-controlled series-compensated transmission line fault classification with discrete wavelet packet entropy measures. Electr. Power Compon. Syst. 42, 554–566 (2014)
- 17. Yang, J.: Identification of green, oolong and black teas in China via wavelet packet entropy and fuzzy support vector machine. Entropy **17**, 6663–6682 (2015)
- 18. Arrais, E., Valentim, R.A.M., Brandao, G.B.: Real time QRS detection based on redundant discrete wavelet transform. IEEE Lat. Am. Trans. 14, 1662–1668 (2016)
- Hamzah, F.A.B., Yoshida, T., Iwahashi, M., Kiya, H.: Adaptive directional lifting structure of three dimensional non-separable discrete wavelet transform for high resolution volumetric data compression. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. E99A, 892–899 (2016)
- Kumar, A., Pooja, R., Singh, G.K.: Performance of different window functions for designing quadrature mirror filter bank using closed form method. Int. J. Signal Imaging Syst. Eng. 8, 367–379 (2015)
- Zhang, Y.D., Dong, Z.C., Ji, G.L., Wang, S.H.: An improved reconstruction method for CS-MRI based on exponential wavelet transform and iterative shrinkage/thresholding algorithm. J. Electromagn. Waves Appl. 28, 2327–2338 (2014)
- Baranwal, N., Singh, N., Nandi, G.C.: Indian sign language gesture recognition using discrete wavelet packet transform. In: International Conference on Signal Propagation and Computer Technology, pp. 573–577. IEEE (2014)
- 23. Gokmen, G.: The defect detection in glass materials by using discrete wavelet packet transform and artificial neural network. J. Vibroengineering **16**, 1434–1443 (2014)
- Qin, Z.J., Wang, N., Gao, Y., Cuthbert, L.: Adaptive threshold for energy detector based on discrete wavelet packet transform. In: Wireless Telecommunications Symposium, pp. 171– 177. IEEE (2012)
- Ghafourian, M., Hassanabadi, H.: Shannon information entropies for the three-dimensional Klein-Gordon problem with the Poschl-Teller potential. J. Korean Phys. Soc. 68, 1267–1271 (2016)
- Phillips, P., Dong, Z., Yang, J.: Pathological brain detection in magnetic resonance imaging scanning by wavelet entropy and hybridization of biogeography-based optimization and particle swarm optimization. Prog. Electromagnet. Res. 152, 41–58 (2015)

- Alcoba, D.R., Torre, A., Lain, L., Massaccesi, G.E., Ona, O.B., Ayers, P.W., Van Raemdonck, M., Bultinck, P., Van Neck, D.: Performance of Shannon-entropy compacted N-electron wave functions for configuration interaction methods. Theor. Chem. Acc. 135 (11), 153 (2016)
- Zhan, T.M., Chen, Y.: Multiple sclerosis detection based on biorthogonal wavelet transform, RBF kernel principal component analysis, and logistic regression. IEEE Access 4, 7567– 7576 (2016)
- 29. Du, S.: Alzheimer's disease detection by Pseudo Zernike moment and linear regression classification. CNS Neurol. Disord. Drug Targets 16, 11–15 (2017)
- Chen, Y.: A feature-free 30-disease pathological brain detection system by linear regression classifier. CNS Neurol. Disord. - Drug Targets 16, 5–10 (2017)
- Lu, S., Qiu, X.: A pathological brain detection system based on extreme learning machine optimized by bat algorithm. CNS Neurol. Disord. - Drug Targets 16, 23–29 (2017)
- Zhou, X.-X.: Comparison of machine learning methods for stationary wavelet entropy-based multiple sclerosis detection: decision tree, k-nearest neighbors, and support vector machine. Simulation 92, 861–871 (2016)
- 33. Zhang, Y.: Binary PSO with mutation operator for feature selection using decision tree applied to spam detection. Knowl.-Based Syst. **64**, 22–31 (2014)
- Abbas, H.A., Belkheiri, M., Zegnini, B.: Feedback linearisation control of an induction machine augmented by single-hidden layer neural networks. Int. J. Control 89, 140–155 (2016)
- 35. Sun, Y.: A multilayer perceptron based smart pathological brain detection system by fractional fourier entropy. J. Med. Syst. 40, 173 (2016)
- Hicham, A., Mohamed, B., Abdellah, E.F.: A model for sales forecasting based on fuzzy clustering and back-propagation neural networks with adaptive learning rate. In: International Conference on Complex Systems, pp. 111–115. IEEE (2012)
- 37. Lu, H.M.: Facial emotion recognition based on biorthogonal wavelet entropy, fuzzy support vector machine, and stratified cross validation. IEEE Access **4**, 8375–8385 (2016)
- Iranmanesh, S.: A differential adaptive learning rate method for back-propagation neural networks. In: Proceedings of the 10th Wseas International Conference on Neural Networks, pp. 30–34. World Scientific And Engineering Acad And Soc (2009)
- 39. Murru, N., Rossini, R.: A Bayesian approach for initialization of weights in backpropagation neural net with application to character recognition. Neurocomputing **193**, 92–105 (2016)
- 40. Li, J.: Detection of left-sided and right-sided hearing loss via fractional fourier transform. Entropy **18**, 194 (2016)
- Kale, M.C., Fleig, J.D., Imal, N.: Assessment of feasibility to use computer aided texture analysis based tool for parametric images of suspicious lesions in DCE-MR mammography, Comput. Math. Method Med. (2013). Article ID: 872676
- Vasta, R., Augimeri, A., Cerasa, A., Nigro, S., Gramigna, V., Nonnis, M., Rocca, F., Zito, G., Quattrone, A.: ADNI: hippocampal subfield atrophies in converted and not-converted mild cognitive impairments patients by a markov random fields algorithm. Curr. Alzheimer Res. 13, 566–574 (2016)
- Hou, X.: Seven-layer deep neural network based on sparse autoencoder for voxelwise detection of cerebral microbleed. Multimedia Tools Appl. (2017). doi:10.1007/s11042-017-4554-8
- 44. Nogueira, R.F., Lotufo, R.D., Machado, R.C.: Fingerprint liveness detection using convolutional neural networks. IEEE Trans. Inf. Forensic Secur. **11**, 1206–1213 (2016)
- 45. Chen, M., Li, Y., Han, L.: Detection of dendritic spines using wavelet-based conditional symmetric analysis and regularized morphological shared-weight neural networks, Comput. Math. Method Med. (2015). Article ID: 454076