Advances in Intelligent Systems Research, volume 143
2nd International Conference on Applied Mathematics, Modelling and Statistics Application (AMMSA 2018)

Four Categories of Human Teeth Based on Biogeography-based Optimization Algorithm and Multilayer Perceptrons

Mengmeng Yang[1*], Arifur Nayeem[2] and Liam O'Donnell[3]
[1]School of Computer Science and Technology, Nanjing Normal University, Nanjing, Jiangsu 210023, P R China
[2]Saidpur government Technical School and College, Saidpur 5310, Bangladesh
[3]School of Engineering, University of Limerick, Limerick, Ireland
[*]Corresponding author

***Abstract*—Teeth is a structure in which many vertebrates exist. For some animals, such as lions, tigers and so on, teeth are chewing tools and weapons to protect themselves. But for human, it also carries the beauty of the face. When the teeth are sick, accurate classification of the teeth seems particularly important. The main purpose of this paper is to classify the teeth accurately using biogeography-based optimization algorithm (BBO) and Multilayer perceptron (MLP). The results showed our method achieved 83.75± 2.95%, 83.50± 5.16%, 84.00± 5.16%, and 84.75± 3.43% accuracy for identifying incisor, canine, premolar, and molar.**

Keywords—teeth classification; biogeography-based optimization algorithm; multilayer perceptrons

I. INTRODUCTION

Teeth is an important organ for human and its function is mainly divided into three aspects [1]:

a) Chewing function: After the food enters the oral cavity, it was cut and torn by the teeth and chemical reaction with the enzyme in the saliva can partially digest the food, which plays an important role in the second digestion of the stomach.

b) Language function: teeth, tongue, lips are the three important organs of pronunciation. The position of the tooth determines the tongue's range of motion at the time of pronunciation, and the positional relationship among the three influences the pronunciation accuracy and clarity.

c) Beautiful appearance: the normal dental relationship of arch and the occlusal can make the cheeks look plump, facial expression natural; if the lack of teeth, then the lip cheeks due to loss of support and collapse, can cause facial aging.

d) Defense function: teeth can also serve as defensive enemies to protect ourselves when we are in danger of being robbed.

It can be seen from the above four points that teeth play a vital role in our diet and appearance [2], so accurately classifying the type of teeth not only can save the time of the patient but also help the doctor started treating the disease earlier when the tooth problems need to be corrected.

There are two types of teeth in our life: deciduous teeth and permanent teeth. Deciduous teeth are the first teeth of a person. Beginning to grow when the baby is six months old and ending about three years old, a total of 20. Permanent teeth are the second and final human teeth. In children six years old, deciduous teeth began to fall off, and permanent teeth began to grow, a total of 32. Because permanent teeth exist for the longest time, this paper mainly classified permanent teeth. According to the morphological characteristics of the teeth can be divided into incisors, canine, premolar, and molars. The first two belong to the front teeth, the latter two belong to the posterior teeth. FIGURE I illustrates the function of these four kinds of teeth.

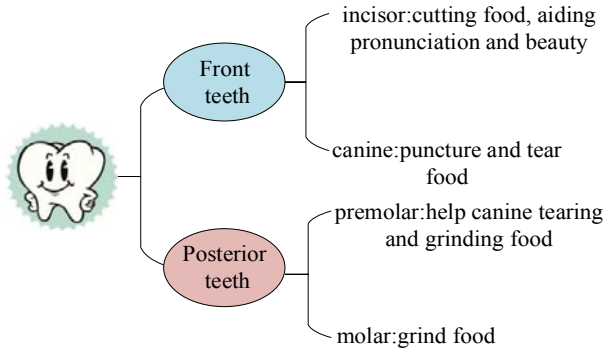

FIGURE I. THE CLASSIFICATION OF TEETH WITH ITS CORRESPONDING FUNCTION

Through the above introduction, we are aware of the importance of teeth classification, while many scholars have also noticed this problem.

Mahoor and Abdel-Mottaleb [3] employed two different Fourier descriptors to extract the features and Bayesian was used as classifier. The author compared the classification performance of the two methods: complex signature and centroid distance, and chose the best method. The t-accuracy of experiment classification is up to 95.5% and the lowest accuracy is 82%. Although the experimental result is good, there is a big problem: just classify the two types of molars and premolar.

Different from the above mentioned in the literature, we propose a novel classification algorithm. First of all, we use wavelet entropy (WE) [4, 5] to extract the features of the tooth images, and input the extracted features into the multilayer

perceptron (MLP) [6, 7]. BBO algorithm [8] is used to train the parameters of MLP to achieve the best performance.

II. MATERIALS

CT scan due to fast scan time, clear images can be clinicians for further treatment design or preoperative choice of path to provide accurate information. Therefore, the experimental dataset selects the CT images of teeth, a total of 160, including 40 images of incisors, canines, premolar and molars, as shown in FIGURE II.

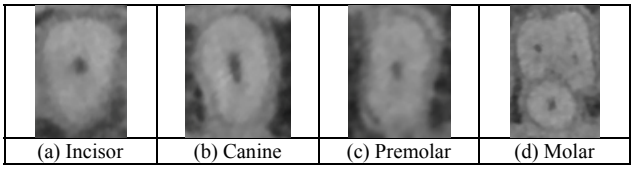

FIGURE II. SAMPLES OF OUR DATASET

III. METHODOLOGY

A. Wavelet Entropy

Wavelet entropy is a term combining wavelet transform [9] and information entropy, which reflects the signal uncertainty and complicated procedure.

Wavelet entropy [10-14] is defined as:

$$WE = -\sum_{i=1}^{n} q_i \log q_i \tag{1}$$

where p_j represents the wavelet energy, the definition of which is as follows

$$\sum_{i=1}^{n} q_i = 1 \quad q_i = \frac{F_i}{F}, \quad F = \sum_{i=1}^{n} F_i \tag{2}$$

where F_i represents the energy component of the signal in each subspace at the same time, F is the total energy at the same time. The step of wavelet entropy extraction feature are as follows [15-18]:

a) Choose wavelet basis;

b) Scale the signal down;

c) Reconstruction of wavelet coefficients;

d) Calculate F_i and F;

e) Calculate WE;

f) Construct eigenvectors;

B. Multilayer Perceptron

MLP is a feed-forward neural network [19-21], it usually consists of three parts: input layer, one or more hidden layers and output layer. Its basic model shown in FIGURE III is as follows (a hidden layer).

The mathematical formula in MLP is as follows:

$$l_k = \sum_{i=1}^{m} W_{ik} x_i + b_k \tag{3}$$

where l_k represents the input of the k-th neuron in the hidden layer. x_1, x_2, x_i, .., x_m is the input unit, W_{ik} represents the connection weight between the i-th neuron in the input layer and the k-th neuron in the hidden layer, b_k is the bias of the k-th hidden node.

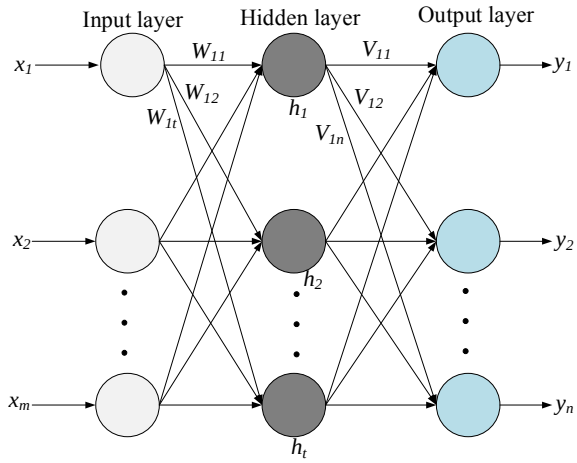

FIGURE III. THE BASIC MODEL OF MLP

The hidden layer output is as follows:

$$r_k = f(l_k) = f_1(\sum_{i=1}^{m} W_{ik} x_i + b_k) \tag{4}$$

where f is the activation function, usually we use sigmoid [22-27], its expression is as follows:

$$f(x) = \frac{1}{1+e^{-x}} \tag{5}$$

Similarly, we can get the input and output of the output layer.

$$q_j = \sum_{k=1}^{t} V_{kj} h_k + d_j \tag{6}$$

where q_j represents the input of the j-th neuron in the output layer, V_{kj} represents the connection weight between the k-th neuron in the hidden layer and the j-th neuron in the output layer, d_j is the bias of the j-th hidden node.

$$y_j = f(q_j) \tag{7}$$

where y_j represents the output of the j-th neuron in the output layer.

From the above formula we can see that the weight and offset values will determine the final output of the output layer. In order to get the actual value of the output closer to the expected

value, it is very important to train the parameters to get the optimal value.

C. Biogeography-Based Optimization Algorithm

BBO is a relatively novel swarm intelligence optimization algorithm proposed by Dan Simon in 2008. BBO algorithm was compared with other optimization algorithm in [28-31], and experiment result showed that the performance of BBO is great. FIGURE IV illustrates the species habitat, each habit contains many characteristic variables such as temperature, precipitation etc. There characteristic variables correspond to the weights and bias in the multilayer perceptron.

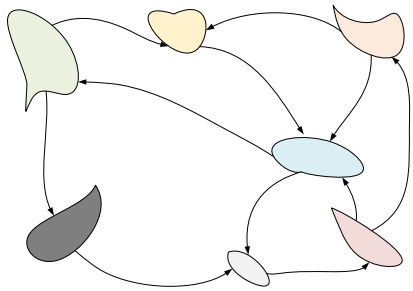

FIGURE IV. HABITAT FOR SPECIES

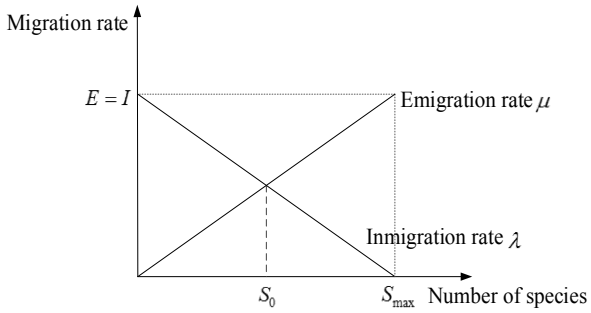

FIGURE V. SPECIES MIGRATION MODEL (E=I)

FIGURE V illustrate the migration of the BBO algorithm. Considering the special case, we set E=I.

IV. Experiments and Results

The 10-fold cross validation was repeated 10 times. The results of our proposed "WE-MLP-BBO" method was listed in TABLE I. The four classes ended up with accuracy of 83.75± 2.95%, 83.50± 5.16%, 84.00± 5.16%, and 84.75± 3.43%.

TABLE I. Cross Validation Results of WE-MLP-BBO Method

	Incisor	Canine	Premolar	Molar
R1	87.50	77.50	80.00	87.50
R2	87.50	87.50	82.50	80.00
R3	82.50	75.00	90.00	82.50
R4	80.00	77.50	85.00	85.00
R5	85.00	87.50	87.50	80.00
R6	87.50	85.00	90.00	87.50
R7	80.00	90.00	85.00	82.50
R8	82.50	87.50	72.50	85.00
R9	82.50	85.00	85.00	90.00
R10	82.50	82.50	82.50	87.50
Average	83.75± 2.95	83.50± 5.16	84.00± 5.16	84.75± 3.43

Next, we compared the proposed BBO algorithm with genetic algorithm (GA) [32] and firefly algorithm (FA) [33]. The results of GA and FA were shown in TABLE II and TABLE III. GA [32] method yielded the accuracy of four teeth classes as 72.50± 5.77%, 72.50± 3.73%, 72.00± 6.95%, and 72.50± 7.17%, respectively. FA [33] method yielded the accuracy of four teeth classes as 76.00± 6.03%, 74.75± 5.58%, 75.50± 5.63%, and 75.75± 5.14%, respectively.

TABLE II. Cross Validation Results of GA Method

	Incisor	Canine	Premolar	Molar
R1	67.50	70.00	77.50	65.00
R2	75.00	77.50	80.00	77.50
R3	65.00	70.00	82.50	80.00
R4	62.50	67.50	65.00	85.00
R5	77.50	75.00	70.00	62.50
R6	70.00	72.50	75.00	70.00
R7	77.50	67.50	70.00	72.50
R8	77.50	72.50	60.00	72.50
R9	75.00	75.00	72.50	75.00
R10	77.50	77.50	67.50	65.00
Average	72.50± 5.77	72.50± 3.73	72.00± 6.95	72.50± 7.17

TABLE III. Cross Validation Results of FA Method

	Incisor	Canine	Premolar	Molar
R1	77.50	77.50	65.00	82.50
R2	65.00	82.50	70.00	82.50
R3	77.50	75.00	75.00	75.00
R4	85.00	80.00	75.00	65.00
R5	82.50	75.00	72.50	72.50
R6	70.00	72.50	80.00	77.50
R7	80.00	67.50	77.50	77.50
R8	72.50	65.00	85.00	77.50
R9	77.50	72.50	80.00	72.50
R10	72.50	80.00	75.00	75.00
Average	76.00± 6.03	74.75± 5.58	75.50± 5.63	75.75± 5.14

It is clear that the proposed BBO is superior to both GA [32] and FA [33]. The reason is BBO contains three efficient strategies: the elite mechanism, the mutation mechanism, and the immigration mechanism. In the future, we shall test some image preprocessing approaches [34-36], to check whether they can increase the performances.

V. Conclusions

This study investigated the application of using biogeography-based optimization in teeth category identification. The results showed the success of proposed method. In the future, we shall try other bioinspired algorithms.

Acknowledgment

The paper is financially supported by Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX17_0342).

References

[1] Aoyama, N., et al., Associations among tooth loss, systemic inflammation and antibody titers to periodontal pathogens in Japanese patients with cardiovascular disease. Journal of Periodontal Research, 2018. 53(1): p. 117-122.

[2] Mastouri, M.H., et al., Genetic study of non-syndromic tooth agenesis through the screening of paired box 9, msh homeobox 1, axin 2, and Wnt

family member 10A genes: a case-series. European Journal of Oral Sciences, 2018. 126(1): p. 24-32.

[3] Mahoor, M.H. and M. Abdel-Mottaleb, Classification and numbering of teeth in dental bitewing images ☆. Pattern Recognition, 2005. 38(4): p. 577-586.

[4] Gorriz, J.M. and J. Ramírez, Wavelet entropy and directed acyclic graph support vector machine for detection of patients with unilateral hearing loss in MRI scanning. Frontiers in Computational Neuroscience, 2016. 10.

[5] Yang, M. Leaf recognition for plant classification based on wavelet entropy and back propagation neural network. in 10th International Conference on Intelligent Robotics and Applications (ICIRA). 2017. Guilin, China: Springer.

[6] Boughrara, H., M. Chtourou, and C.B. Amar, MLP neural network using constructive training algorithm: application to face recognition and facial expression recognition. International Journal of Intelligent Systems Technologies & Applications, 2017. 16(1): p. 53.

[7] Khundrakpam, J., et al., MLP-GA based algorithm to detect application layer DDoS attack. Journal of Information Security & Applications, 2017. 36: p. 145-153.

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

[9] Yang, M., Dual-Tree Complex Wavelet Transform and Twin Support Vector Machine for Pathological Brain Detection. Applied Sciences, 2016. 6(6).

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

[11] Phillips, P., Z. Dong, and J. Yang, Pathological brain detection in magnetic resonance imaging scanning by wavelet entropy and hybridization of biogeography-based optimization and particle swarm optimization. Progress In Electromagnetics Research, 2015. 152: p. 41-58.

[12] Sun, P., Pathological brain detection based on wavelet entropy and Hu moment invariants. Bio-Medical Materials and Engineering, 2015. 26(s1): p. 1283-1290.

[13] 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, 2016. 92(9): p. 861-871.

[14] 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.

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

[16] Wang, S.-H. Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. Multimedia Tools and Applications, 2016. DOI: 10.1007/s11042-016-4222-4.

[17] Gorriz, J.M. Multivariate approach for Alzheimer's disease detection using stationary wavelet entropy and predator-prey particle swarm optimization. Journal of Alzheimer's Disease, 2017. DOI: 10.3233/JAD-170069.

[18] Han, L., Identification of Alcoholism based on wavelet Renyi entropy and three-segment encoded Jaya algorithm. Complexity, 2018. 2018.

[19] Fabijańska, A., A. Fabijańska, and A. Fabijańska. Corneal Endothelium Image Segmentation Using Feedforward Neural Network. in Federated Conference on Computer Science and Information Systems. 2017.

[20] Foody, G., Impacts of Sample Design for Validation Data on the Accuracy of Feedforward Neural Network Classification. Applied Sciences, 2017. 7(9): p. 888.

[21] Wang, S., et al., Fruit Classification by Wavelet-Entropy and Feedforward Neural Network Trained by Fitness-Scaled Chaotic ABC and Biogeography-Based Optimization. Entropy, 2015. 17(8): p. 5711-5728.

[22] Ji, G., Fruit classification using computer vision and feedforward neural network. Journal of Food Engineering, 2014. 143: p. 167-177.

[23] Wu, L., Weights optimization of neural network via improved BCO approach. Progress in Electromagnetics Research, 2008. 83: p. 185-198.

[24] Zhang, Y., Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. Expert systems with applications, 2009. 36(5): p. 8849-8854.

[25] Wu, L. and G. Wei, A New Classifier for Polarimetric SAR Images. Progress in Electromagnetics Research, 2009. 94: p. 83-104.

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

[27] Huo, Y., Feature Extraction of Brain MRI by Stationary Wavelet Transform and its Applications. Journal of Biological Systems, 2010. 18(S): p. 115-132.

[28] Li, N., Conversion prediction of mild cognitive impairment to Alzheimer's disease via hybridization of two intelligent algorithms: biogeography-based optimization and particle swarm optimization. Basic & Clinical Pharmacology & Toxicology, 2016. 119(S2): p. 9.

[29] Wu, X., Smart detection on abnormal breasts in digital mammography based on contrast-limited adaptive histogram equalization and chaotic adaptive real-coded biogeography-based optimization. Simulation, 2016. 92(9): p. 873-885.

[30] Wu, J., Fruit classification by biogeography-based optimization and feedforward neural network. Expert Systems, 2016. 33(3): p. 239-253.

[31] Du, S., Multi-objective path finding in stochastic networks using a biogeography-based optimization method. Simulation, 2016. 92(7): p. 637-647.

[32] Chen, R.H., et al., Prediction of LBB leakage for various conditions by genetic neural network and genetic algorithms. Nuclear Engineering and Design, 2017. 325: p. 33-43.

[33] Sanchez, D., P. Melin, and O. Castillo, Optimization of modular granular neural networks using a firefly algorithm for human recognition. Engineering Applications of Artificial Intelligence, 2017. 64: p. 172-186.

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

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

[36] Wu, L., A hybrid method for MRI brain image classification. Expert Systems with Applications, 2011. 38(8): p. 10049-10053.