EAI Endorsed Transactions
on e-Learning

Research Article EAI.EU

# Hearing loss classification via stationary wavelet entropy and Biogeography-based optimization

Chong Yao[1,*], Chaosheng Tang[1], Junding Sun[1]

[1]Henan Polytechnic University, Jiaozuo, Henan, China

## Abstract

**Introduction:** Sensorineural hearing loss is associated with many complications and needs timely detection and diagnosis.
**Objectives:** Optimize the sensorineural hearing loss detection system to improve the accuracies of image detection.
**Method:** The stationary wavelet entropy was used to extract the features of NMR images, the single hidden layer neural network was used for classification, and the BBO algorithm was used for optimization to avoid the dilemma of local optimum. We used two-level SWE as input to the classifier to enhance the identify and classify ability of hearing loss.
**Results:** The results of 10-fold cross validation show that the accuracies of HC, LHL and RHL are 91.83± 3.09%, 92.67±2.38% and 91.17±2.61%, respectively. The overall accuracy is 91.89±0.70%.
**Conclusion:** This model has good performance in detecting hearing loss.

**Keywords:** Virtual Private Network, designing secure enterprise network, secure enterprise network.

Received on 23 June 2020, accepted on 30 July 2020, published on 07 August 2020



doi: 10.4108/eai.7-8-2020.165964

## 1. Introduction

Sensorineural hearing loss (SNHL) is the most common sensory deficit in the world, which caused by the dysfunction of one or more parts of the auditory pathway between the inner ear and the auditory cortex [1]. It is primarily manifested by unilateral or bilateral ears progressive hearing impairment at different levels even deafness, accompanied with tinnitus, sensation of intra-aural occlusion and the like [2]. It often leads to depression, falls, lower intelligence, speech and language delay, and other complications.

SNHL brings language, communication and even psychological barriers to patients, seriously affects the work and life, and brings huge social and economic burden to the country. MRI can show the lesions of soft tissue and intracranial structure well, especially the application of MRI water imaging technology makes it possible to display the fine structure of the inner eardrum labyrinth. Therefore, MRI can be used to identify brain atrophy as the evaluation

*Corresponding author. Email: Yaochong@home.hpu.edu.cn

standard for neurological hearing loss, so as to achieve the purpose of detection [3].

Therefore, based on the above characteristics, scholars generally choose to use computer vision combined with machine learning to detect hearing loss [4–8], and previous experiments have also shown the effectiveness of this system. But many existing detection systems easily fall into the dilemma of local optimization when training and optimizing neural networks. We put forward a new idea: we used the BBO algorithm to replace the previous optimization algorithm, based on its genetic variation characteristics, to avoid the short-term blind optimal state of the system. Besides, we proposed to use SWE to avoid large deviation of experimental results due to small changes. Finally, the overall accuracy of our model reached 91.89±0.70%.

## 2. Background

Several methods have been proposed to detect hearing loss images and brain images in the past. O.Profant (2014) et al [9] used MR morphometry and DTI to study SNHL. They mainly study the physiological changes of hearing loss patients, which makes a great contribution to the image analysis research of later scholars. F. Liu(2017) et al [10] proposed to combine wavelet entropy with feedforward neural network trained by genetic algorithm to defect hearing loss. Their method using 4-level decomposition yielded the overall accuracy of 81.11±1.34%. Their earlier application of genetic algorithm for hearing loss detection system provides ideas for many subsequent researchers. Fang-zhou BAO (2018) et al [11] defected hearing loss via Wavelet Entropy and Particle Swarm Optimized Trained Support Vector Machine. They chose wavelet entropy to extract the features of the image and chose two-level decomposition in the calculation. The final accuracy results of HC, LHL, RHL and overall were 85.20±3.79%, 85.20±4.64%, 86.40±5.06 % and 85.60±0.84% respectively.

In the process of referring to previous studies, we found there are common problems in defecting hearing loss:

(i) How to extract more information from MRI images?

(ii) How to select a classifier to make the results more robust?

(iii) How to optimize after selecting classifier?

These questions are raised to ensure that we can get better performance in hearing loss detection. Based on the above problems, we propose the following solutions in this paper: our team used SWE to extract features and proposed using BBO algorithm, which can have a global optimization to the neural network and the image is classified by single-hidden-layer neural network. We will explain the model we chose in the next sections of the article.

The remaining parts of this paper are organized as follows: Section 2 shows the background of the hearing loss detection. Section 2 provides the date source. Section 1 introduces the basic principle of SWE, the construction of single-hidden-layer neural network and the optimization principle of BBO algorithm. Section 5 introduces the experimental results and data analysis. Section 1 gives the conclusion.

Table 1. The data analysis of all subjects.

| | Age (year) | Gender (m/f) | Disease duration (year) | PTA of left ear (dB) | PTA of right ear (dB) | Education level (year) |
|---|---|---|---|---|---|---|
| Control | 53.2± 5.9 | 27/33 | 0 | 23.1± 2.1 | 20.3± 2.2 | 11.7± 3.2 |
| RHL | 51.3± 8.3 | 29/31 | 13.8±14.5 | 22.0± 3.5 | 80.4±18.8 | 12.3± 2.5 |
| LHL | 51.4± 9.0 | 32/28 | 17.1±18.2 | 77.3±17.2 | 20.1± 4.5 | 12.8± 1.5 |

(PTA= pure tone average)

## 3. Dataset

Our team selected 180 subjects: 60 patients with right-sided hearing loss (RHL), 60 patients with left-sided hearing loss (LHL), and 60 healthy controls (HC). Due to ethical issues and the cost of MRI scans, it is difficult to obtain a large sample of patients with hearing loss, but in a similar number of sample sets, many researchers have done studies and obtained good performance. In order to reduce the influence of other diseases on our model, we excluded subjects with known mental or neurological diseases, brain injury (such as tumor or stroke), psychotropic medication, and contraindications to magnetic resonance imaging from the sample selection. The epidemiological investigation and analysis results of the patients are shown in the Table 1.

This part mainly investigates the gender, age distribution and other aspects of the disease. The normal features of the inner ear and image detection standard are shown as Table 2. MRI system settings and Pure tone average (PTA) evaluation criteria are consistent with other experiments, using a 3.0-T MRI system and six different octave frequencies (250, 500, 1000, 2000, 4000 and 8000 Hz).

Table 2. Inner ear detection standard.

| | Normal | Abnormal |
|---|---|---|
| Cochlea | $2_{1/2}$～$2_{3/4}$ circles<br>30 ~ 32mm（total length)<br>5mm( from the bottom to the top） | Incomplete segmentation, all levels are ambiguous. |
| Vestibular | 5mm × 3mm × 5mm(Long × wide × high） | Vestibular expand（wide > 3.2mm） |
| Vestibular aqueduct | ≤1. 5mm（width at the midpoint between the main foot and the outer mouth） | Vestibular aqueduct expand (> 1.5mm or outer orifice width >2mm) |
| Inner Ear Routine | 4~6mm(width) | narrow（width<4mm)<br>expand（width>6mm) |

# 4. Methodology

## 4.1. Stationary Wavelet Entropy

Scholars have used many computer-aided diagnosis (CAD) systems to replace the difficult method of manual labeling in feature extraction these years. Bansal *et al* (2020) [12] used the Bag of Features method to represents the images as an order less collections of local features. Nha *et al* (2015) [13] used the Stationary wavelet transform (SWT) to predicts gene expression. Discrete wavelet transform (DWT) and wavelet entropy (WE) are more common method but exist a problem of preserving the translation invariance property.

In response to this question, wang (2018) *et al* [14] introduced a novel feature named as stationary wavelet entropy (SWE), which combined stationary wavelet transform (SWT) and Shannon entropy. Three different numbers of features were used in the experiment, but the results all showed that the accuracy of SWE was significantly higher than that of other wavelets, nearly 100%. Therefore, compared with three traditional features that were successfully applied in pathological brain detection, SWE has better classification ability.

The one-level SWE utilize filters to decompose image and extract the entropy of the four sub-bands with names of $LL_1$, $LH_1$ and $HH_1$ as shown in Figure 1.

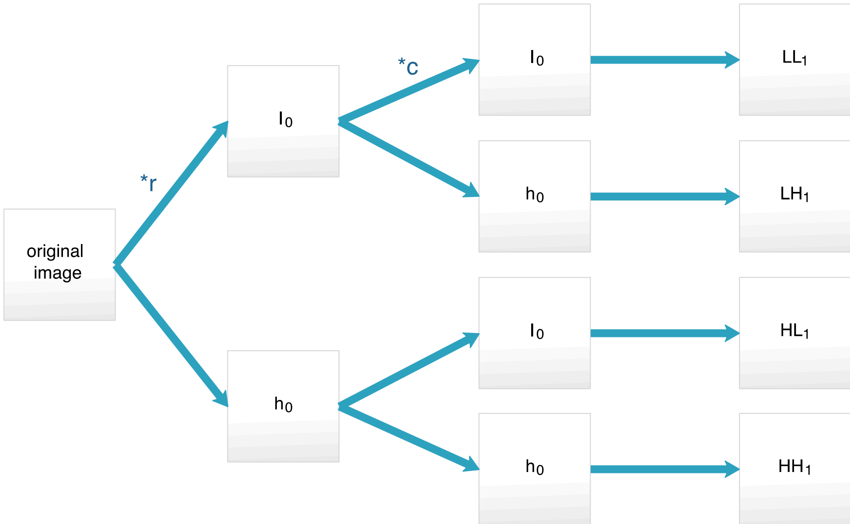


**Figure 1.** Illustration of Level-1 of Stationary Wavelet Entropy.
( $*_r$ represents row-wise filter and $*_c$ represents column-wise filter, l and h represent the low-pass and the high-pass filters respectively)

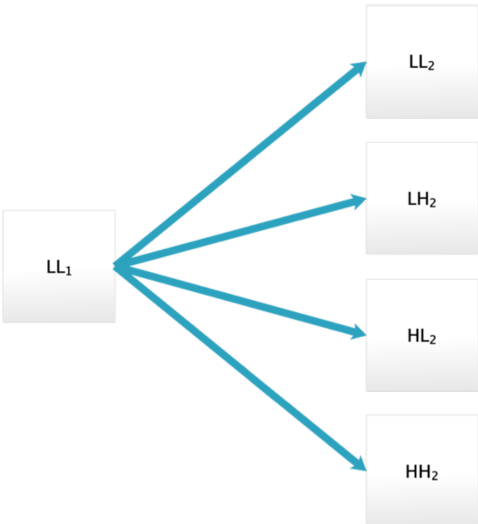


**Figure 2.** Illustration of Level-2 of Stationary Wavelet Entropy.

The two-level SWE use $LL_1$ sub-band to perform another one-level SWE as shown in Figure 2. And so on, the four sub-bands of different levels are shown below:

$$LL_{k+1} = (LL_k) *_r (h_k) *_c (h_k) \quad (1.)$$
$$LH_{k+1} = (LL_k) *_r (h_k) *_c (g_k) \quad (2.)$$
$$HL_{k+1} = (LL_k) *_r (g_k) *_c (h_k) \quad (3.)$$
$$HH_{k+1} = (LL_k) *_r (g_k) *_c (g_k) \quad (4.)$$

Then, we calculate the entropy values. We choose Shannon entropy with definition of

$$\text{Entropy} = -\sum_i X_i^2 log X_i^2 \quad (5.)$$

Here $X_i$ represents the $i$-th element of a given subband. Combined with the above, the SWE pseudocode is listed in the Table 3.

Table 3. Algorithm of Stationary Wavelet Entropy.

| Stationary Wavelet Entropy (SWE) |
|---|

| | |
|---|---|
| Step A | Import the preprocessed MRI images; |
| Step B | Select the best wavelet in the wavelet family; |
| Step C | Choose decomposition level $m$; |
| Step D | Perform stationary wavelet transform (SWT) on the given images; |
| Step E | Generate and record $(3m + 1)$ wavelet subbands; |
| Step F | Use formula 5 to calculate entropy over each subband; |
| Step G | Vectorize all the entropy results and output it as the feature to input layer |

## 4.2. Single-hidden-layer feedforward neural network

There are many efficient classifiers, which all show high accuracy in different fields: such as Backpropagation Neural Network used for cervical cancer classification [15], SVM utilized for numerous real-world applications/problems [16], deep learning used for diseased pinus trees recognition [17]. And in recent years, many neural network architectures based on the bionics have been proposed, while multi-layer feedforward neural network (FFNN) is considered as one of fundamental architectures. Deep learning based approach will also be the trend of future research [18]. But it is very time-consuming to train when FFNN is used for classification or regression [19]. Hence, single-hidden-layer feedforward neural network (SLFN) becomes an optimal selection [20].

SLFN includes input layer, output layer and hidden layer. There is no feedback in the whole network, that is, directed acyclic graph. Its application in this experiment is shown as Figure 3. the input layer contains seven neurons, since we take a 2-level SWE including seven features. We use two neurons in the output layer to represent three types of results: RHL, LHL and HC. No matter how many neurons there are in the input layer and output layer, as long as the number of neurons in the hidden layer is reasonable or sufficient, it can approximate any function.

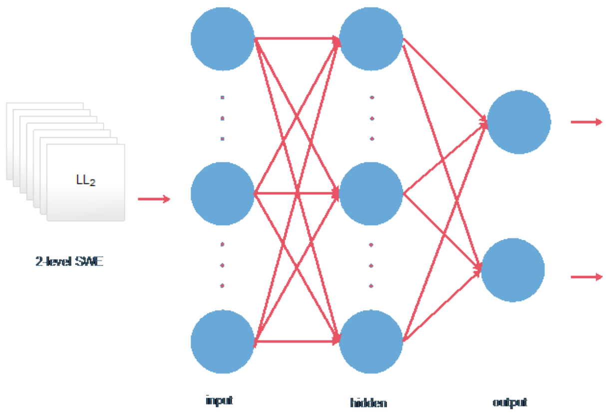


**Figure 3.** Diagram of the SLFN.

Deep learning approaches [21−29] were not used since our dataset is of small size, the deep neural networks require a large dataset. After the input layer receives the image dataset from SWE. The learning function can be defined as

$$F(x) = \sum_{i=1}^{L} \beta_i G(w_i x + b_i) \tag{6.}$$

where $i$ is the $i$ th hidden node, $w_i$ is inner weight connecting the input layer with the $i$th hidden node, $\beta_i$ is outer weight connecting the $i$th hidden node with the output node, $b_i$ is the value of the $i$th hidden node. $G$ is active function, which should be continuous. $L$ is the total number of the neurons.

The most important thing in machine learning is the selection of function and the optimization of weight parameters.

We can chose select function according to the dataset, such as

$$G(x) = \frac{1}{1 + e^{x^2}} \tag{7.}$$

And selecting the optimal parameter can be translated into the following formula

$$|F(x) - f(x)| < \varepsilon \qquad (8.)$$

where $f(x) = y_j$ , $j = 1,2,3, \ldots, N$.

In order to find the optimal parameter, we just need to find the parameter that minimizes $\varepsilon$.

## 4.3. Biogeography-based optimization

Many optimization algorithms have been proposed in recent years, such as Particle Swarm Optimization (PSO) [30], ant colony optimization (ACO), Shuffled Frog Leaping Algorithm (SFLA),Harmony Search (HS) [31] and Biogeography-based optimization(BBO) and so on. Biogeography-based Optimization (BBO) algorithm has been widely used since it was proposed by Simon(2008) [32] because of its advantages such as low problem dependence, few algorithm parameters and easy implementation. X. Zhang et al (2020 ) [33] used improved Laplacian Biogeography-Based Optimization Algorithm for Quadratic Assignment Problems (QAPs) and Q. Niu et al(2014) [34] used BBO for model parameter estimation of solar and fuel cells. They all performed high accuracy rate.

In our paper we select the BBO algorithm, because of its strong mining capacity, integer coding, less time, fast convergence and not easy to fall into the local optimal [35].

In Biogeography-based optimization algorithm, each solution is considered as a habitat with a habitat suitability index (HSI). The factors that affect HSI are called Suitable Index Vector (SIV) like climate. Each habitat has its own immigration rate ($\lambda$)and emigration rate ($\mu$). Habitats with a high HIS have a low species immigration rate and a high emigration rate. On the contrary, a low HIS means a high species immigration rate and a low emigration rate. According to the migration of species between different habitats, the habitat with low HIS value can obtain more information form the habitat with high HIS value, so as to realize the continuous evolution of the habitat. Our team mainly use migration and mutation models of species to optimize the neural network.

In BBO algorithm, immigration rate and emigration rate are expressed as

$$\boldsymbol{\lambda_{i=}I\left(1 - \frac{S_i}{S_{max}}\right)} \qquad (9.)$$

$$\boldsymbol{\mu_{i=}E\left(\frac{S_i}{S_{max}}\right)} \qquad (10.)$$

where $\boldsymbol{I}$ is the biggest immigration rate, E is the largest emigration rate, $S_i$ is the current population quantity, $S_{max}$is the maximum population size. The relationship can be shown in Figure 4.

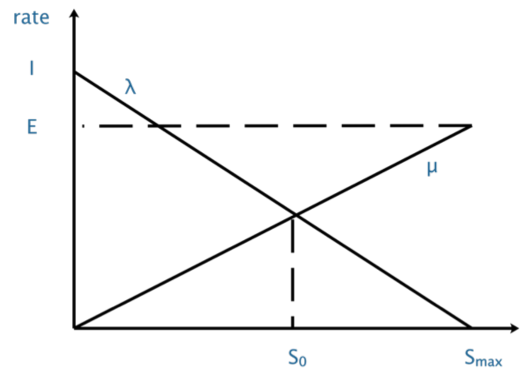


**Figure 4.** Relationship plot of species number and migration rate.

Suppose the maximum possible migration rate of this habitat is I and take the maximum when the species of this habitat is zero. As the number of species increases, habitats become more crowded, fewer species are able to successfully move into the habitat, and the rate of migration is decreasing. The maximum number of species that can be sustained in the habitat is 0.

In addition to migration, habitat can also be changed by sudden disaster, which is called mutation in biogeography. The probability of variation $m_i$ in a habitat with probability $P_i$ can be written as

$$m_i = m_{max}\left(1 - \frac{P_i}{P_{max}}\right) \qquad (11.)$$

$m_{max}$ is a maximum variation rate determined according to the actual situation of the optimization problem. $P_{max}$ is the maximum rate of species existence probability. Migration operation can be described in Algorithm 1, and mutation operation can be described in Algorithm 2:

Algorithm 1. Pseudocode of migration.

| |
|---|
| Select $H_i$ with probability $\propto$ $\lambda_i$ |

If $H_i$ is selected
For j=1 to n
Select $H_j$ with probability $\propto \mu_i$
If $H_j$ is selected
Randomly select an SIV $\sigma$ from $H_j$
Replace a random SIV in $H_i$ $with\ \sigma$
End
End
End
($H_i$ $and$ $H_j$ represent habitat.)

Algorithm 2. Pseudocode of mutation.

For j=1 to m
Use $\lambda_i$ and $\mu_i$ to compute the probability $P_i$
Select SIV $H_{i(j)}$ with probability $\propto P_i$
If $H_{i(j)}$ is selected
Replace $H_{i(j)}$ with a randomly generated SIV
End
End

The pipeline of BBO algorithm can be shown in Figure 5. As with other optimization algorithms, BBO also incorporate some sort of elitism in order to retain the best solutions.

## 4.4. Implementation

Overall, the identification system we purposed to defect can be depicted in Figure 6. BBO algorithm take the sample error as the target function. The flow of single-hidden-layer neural network optimization based on BBO algorithm is in Algorithm 3.

## 4.5. Measure

In order to avoid the overfitting phenomenon in the experiment, we adopt the cross validation. To be more precise, we take 10 runs 10-fold cross validation. For each fold, there are 6 LHL images, 6 RHL images and 6 HC images. Eight folds are used for training, one fold is used for validation and the final one fold is used for test in each trial.

The ideal confusion matrix of a classifier over a $10 \times 10$-fold cross validation takes following form as:

$$\mathrm{E}(r = 10, f = 10) = \begin{bmatrix} 600 & 0 & 0 \\ 0 & 600 & 0 \\ 0 & 0 & 600 \end{bmatrix} \quad (12.)$$

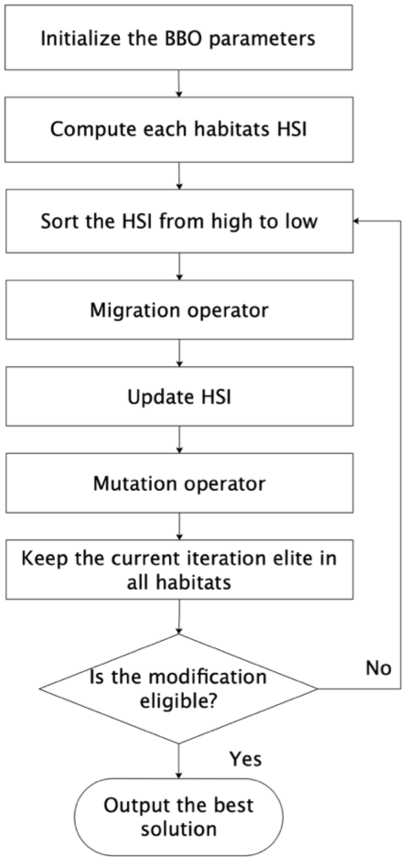


**Figure 5.** Pipeline of BBO algorithm.

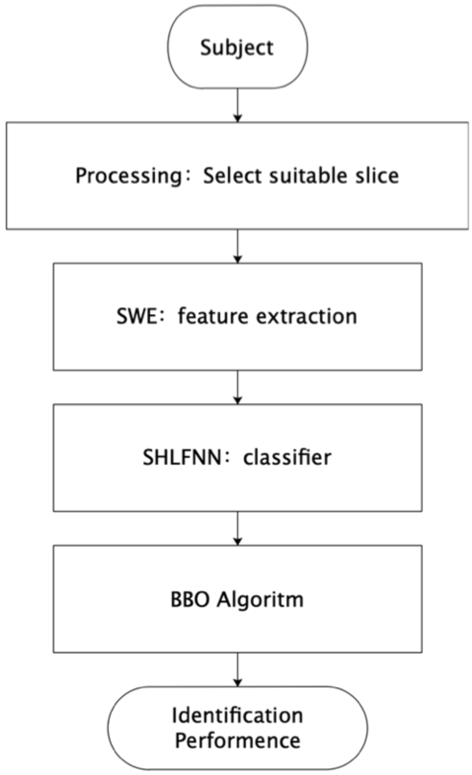


**Figure 6.** Block diagram of the whole system.

Algorithm 3. Pseudocode of BBO.

1. According to the initialization process, randomly generate the initial habitat group which is composed of SIV. The population size N, habitat immigration rate $\lambda$, emigration rate $\mu$, mutation $m$, was initially set.
2. Map each SIV vector to a set of weights for the network.
3. A certain number of samples are randomly selected from the sample space to form training samples for training.
4. Calculate the HSI for each habitat and sort it.
5. If the smallest HSI is smaller than the given number, stop the operation and output the weight as the optimization result. If not, keep going.
6. Replace and update a vector based on migration and mutation operations. Go on to step 4.

Here r represents run number and f the fold number. Column and column vectors are HC, LHL, RHL. The overall accuracy (OA) and the sensitivity of $\kappa$th class $s(k)$ are commonly used in classification problems. They are defined as

$$OA = \frac{\sum_{i=1}^{3} E_{ii}(r = 10, f = 10)}{\sum_{i=1}^{3}\sum_{j=1}^{3} E_{ij}(r = 10, f = 10)} \quad (13.)$$

$$s(k) = \frac{E_{kk}(r = 10, f = 10)}{\sum_{i=1}^{3} E_{ki}(r = 10, f = 10)} \quad (14.)$$

Here $r$ and $f$ are defined in the same way as described above. They are all represent 10 runs and 10 folds. In other words, $OA$ can be understood as the accuracy of the experiment as a whole, which can be obtained by dividing diagonal elements with all elements, while $s$ is the accuracy of a certain kind, which can be obtained by dividing diagonal elements with the whole elements in the row.

# 5. Experiment Results and Discussions

## 5.1. Analysis of Our Method SWE-BBO

In this experiment, we used two-level decomposition and db4 wavelet to extract features. Moreover, BBO algorithm is proposed to optimize the neural network classifier. We will give the reasons one by one in the following sections. First, analyze the overall experimental results. The sensitivity and overall accuracy (OA) of 10-fold stratified cross validation are listed in Table 4. The sensitivities over the three subject classes are 91.83± 3.09%, 92.67± 2.38% and 91.17± 2.61%, respectively. The overall accuracy is 91.89± 0.70% high with a small error.

Table 4. Performance analysis of SWE-BBO.

| Fold | S1 | S2 | S3 | OA |
|---|---|---|---|---|
| F1 | 90.00 | 91.67 | 93.33 | 91.67 |
| F2 | 95.00 | 91.67 | 88.33 | 91.67 |
| F3 | 91.67 | 93.33 | 93.33 | 92.78 |
| F4 | 90.00 | 93.33 | 90.00 | 91.11 |
| F5 | 95.00 | 90.00 | 90.00 | 91.67 |
| F6 | 91.67 | 93.33 | 90.00 | 91.67 |
| F7 | 85.00 | 96.67 | 95.00 | 92.22 |
| F8 | 91.67 | 95.00 | 86.67 | 91.11 |
| F9 | 93.33 | 93.33 | 93.33 | 93.33 |
| F10 | 95.00 | 88.33 | 91.67 | 91.67 |
| Mean+SD | 91.83± 3.09 | 92.67± 2.38 | 91.17± 2.61 | 91.89± 0.70 |

(S1=LHL, S2=RHL, S3=HC)

In order to make the data more vividly represented, we can see the line chart of this experiment in Figure7. It is obvious from the figure that the accuracies of this test are all above 88%, and the overall accuracy is stable at around 91%, which means that our method is very robust. These data are sufficient to confirm the good performance of our model. As a very rigorous discipline, medical science cannot ignore the mistakes of individual sample while pursuing a high degree of overall accuracy. A small mistake may affect the life of patients.

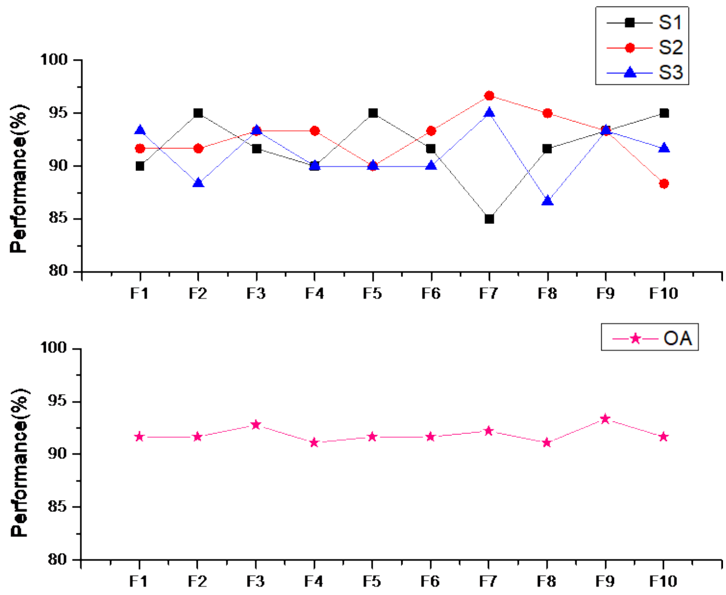


**Figure 7.** Line chart of performance of SWE-BBO.

We can see that the accuracy of different types of images are not different, and they all show good performance. But it also illustrates that the right hearing loss is easier to identify. This may be due to the significant increase in ALFF and fALFF values in patients with unilateral SSNHL during the acute phase. It suggests that the resting brain function of patients with unilateral SSNHL may be hyperactivated during the acute period. It has also been suggested that the level of activity in the right medial temporal gyrus could be used as a potential imaging biomarker to assess the degree of auditory impairment.

## 5.2. Comparison of WE and SWE

As we mentioned in the previous section, many experiments used traditional WE before. Therefore, we tested using the WE and BBO algorithms to get better persuasion. Table 5 shows us the sensitivity and overall accuracy results of WE-BBO under the same conditions as SWE-BBO we mentioned before. The sensitivities over the three subject classes are 85.17± 2.28%, 87.17± 2.23% and 86.00± 3.53%. The overall accuracy is 86.11± 1.08%. In addition, in this section, we will explain why we chose SWE.

Table 5. Performance analysis of WE-BBO.

| Fold | S1 | S2 | S3 | OA |
|---|---|---|---|---|
| F1 | 81.67 | 90.00 | 91.67 | 87.78 |
| F2 | 86.67 | 88.33 | 83.33 | 86.11 |
| F3 | 83.33 | 88.33 | 81.67 | 84.44 |
| F4 | 85.00 | 86.67 | 88.33 | 86.67 |
| F5 | 86.67 | 90.00 | 80.00 | 85.56 |
| F6 | 83.33 | 88.33 | 86.67 | 86.11 |
| F7 | 85.00 | 85.00 | 88.33 | 86.11 |
| F8 | 88.33 | 85.00 | 88.33 | 87.22 |
| F9 | 83.33 | 83.33 | 86.67 | 84.44 |
| F10 | 88.33 | 86.67 | 85.00 | 86.67 |
| Mean+SD | 85.17± 2.28 | 87.17± 2.23 | 86.00± 3.53 | 86.11± 1.08 |

Compared the WE-BBO with SWE-BBO, it's easy to see a significant difference in accuracy. Replacing WE with SWE can have an increase from 86.11± 1.08% to 91.89± 0.70% in overall accuracy. In order to make the experimental results clearer, we made error bar of the two methods in Figure 8. Each indicator of SWE-BBO in the figure is higher than WE-BBO and occupies the upper part of the y-value range. Therefore, we chose the one with better performance. This may be because SWT, which makes up SWE, provides more information than WT, which makes up WE.

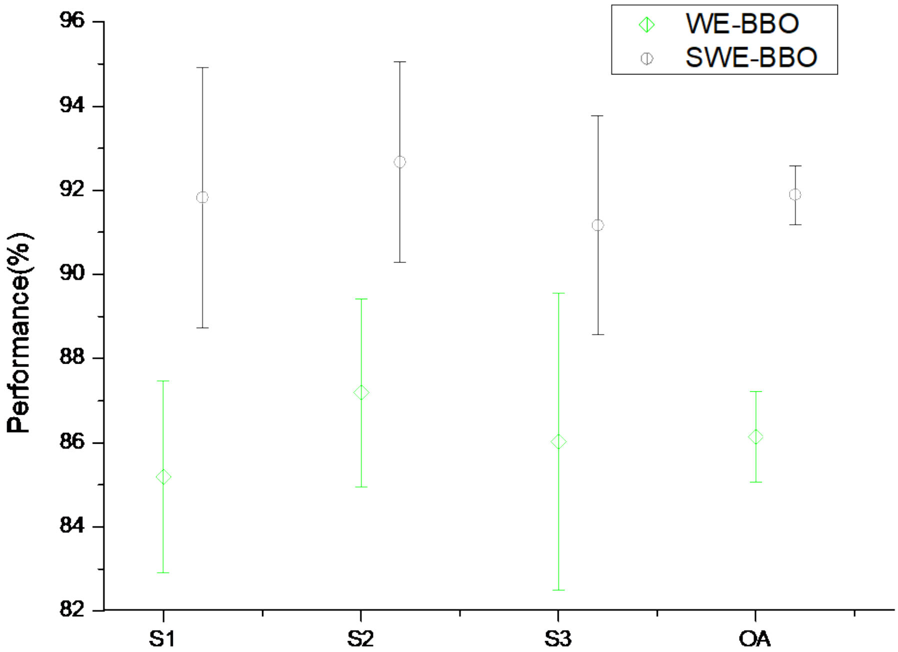


**Figure 8.** Error bar of WE-BBO and SWE-BBO.

## 5.3. Optimal wavelet selection

In this experiment, we fix the decomposition level as 2, which reason will be given in the following section, and choose the Daubechies wavelet which is the most popular among the wavelet family as db4. In this section we will do some experiments with varying the wavelets to see why our choice was optimal. The experimental results of db1, db2 and db3 can be shown in Table 6. We compared the data in this section with the experimental results in Table 4 (db4) and get Figure 9, which significantly shows the superiority of the db4 we chose.

Table 6. Performance analysis of using other wavelets.

| **db1** | **Sensitivity** | | | **Overall accuracy** |
|---|---|---|---|---|
| Fold | S1 | S2 | S3 | |
| F1 | 86.67 | 86.67 | 86.67 | 86.67 |
| F2 | 88.33 | 85.00 | 86.67 | 86.67 |
| F3 | 83.33 | 85.00 | 85.00 | 84.44 |
| F4 | 86.67 | 83.33 | 88.33 | 86.11 |
| F5 | 76.67 | 86.67 | 86.67 | 83.33 |
| F6 | 88.33 | 90.00 | 83.33 | 87.22 |
| F7 | 86.67 | 80.00 | 88.33 | 85.00 |
| F8 | 80.00 | 83.33 | 91.67 | 85.00 |
| F9 | 88.33 | 85.00 | 85.00 | 86.11 |

| | | | | |
|---|---|---|---|---|
| F10 | 88.33 | 83.33 | 83.33 | 85.00 |
| Mean+SD | 85.33±4.07 | 84.83±2.66 | 86.50±2.54 | 85.56±1.20 |
| **db2** | **Sensitivity** | | | **Overall accuracy** |
| Fold | S1 | S2 | S3 | |
| F1 | 93.33 | 86.67 | 85.00 | 88.33 |
| F2 | 85.00 | 90.00 | 91.67 | 88.89 |
| F3 | 85.00 | 88.33 | 90.00 | 87.78 |
| F4 | 83.33 | 90.00 | 88.33 | 87.22 |
| F5 | 88.33 | 93.33 | 88.33 | 90.00 |
| F6 | 88.33 | 90.00 | 85.00 | 87.78 |
| F7 | 86.67 | 90.00 | 88.33 | 88.33 |
| F8 | 86.67 | 86.67 | 86.67 | 86.67 |
| F9 | 91.67 | 90.00 | 91.67 | 91.11 |
| F10 | 83.33 | 90.00 | 86.67 | 86.67 |
| Mean+SD | 87.17±3.34 | 89.50±1.93 | 88.17±2.42 | 88.28±1.42 |
| **db3** | **Sensitivity** | | | **Overall accuracy** |
| Fold | S1 | S2 | S3 | |
| F1 | 85.00 | 90.00 | 93.33 | 89.44 |
| F2 | 90.00 | 90.00 | 85.00 | 88.33 |
| F3 | 90.00 | 86.67 | 91.67 | 89.44 |
| F4 | 86.67 | 91.67 | 86.67 | 88.33 |
| F5 | 91.67 | 91.67 | 83.33 | 88.89 |
| F6 | 91.67 | 90.00 | 88.33 | 90.00 |
| F7 | 88.33 | 91.67 | 90.00 | 90.00 |
| F8 | 88.33 | 86.67 | 91.67 | 88.89 |
| F9 | 88.33 | 90.00 | 90.00 | 89.44 |
| F10 | 93.33 | 86.67 | 90.00 | 90.00 |
| Mean+SD | 89.33±2.51 | 89.50±2.09 | 89.00±3.16 | 89.28±0.64 |

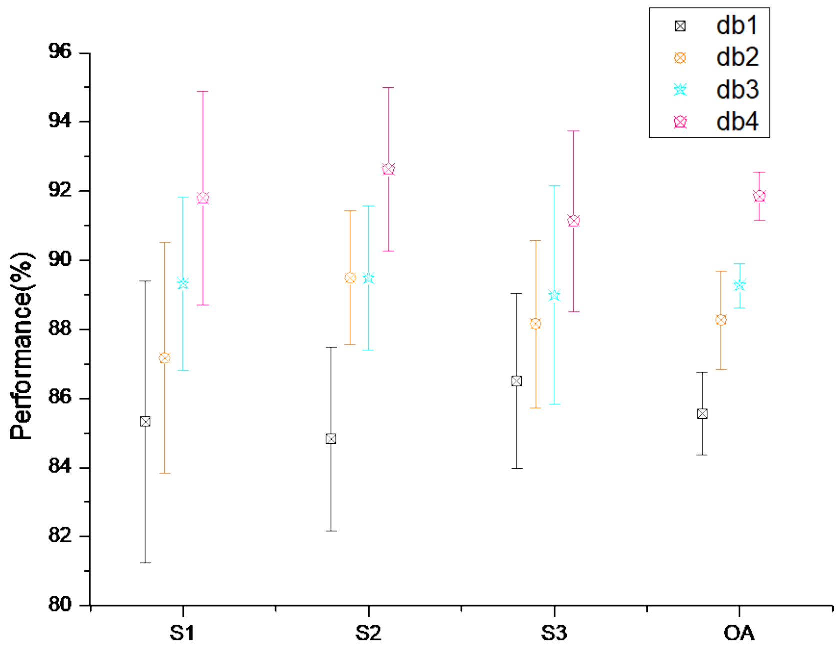


**Figure 9.** Error bar of different wavelet experiment.

The error bar of four different wavelet experiments show that the accuracy of db4 is the highest in each class of detection. The overall accuracy of db4 with 91.89±0.70% was higher than that of db1 with 85.56±1.20%, db2 with 88.28±1.42% and db3 with 89.28±0.64%. This is enough to confirm the choice of our experiment is correct. At the same time, it is not difficult to find that with the change of wavelet, the accuracy of experimental detection is gradually improved, and the performance is constantly optimized.

## 5.4. Optimal Decomposition Level

In this section, we will present the experimental process in which we choose the 2-level decomposition. In order to find the optimal decomposition level, we tested the accuracy and sensitivity of three subjects at several different decomposition levels. Suppose there are four levels of decomposition with the values of 1-4. The final results are listed in Table 7 and Figure 10, and their overall accuracy are 89.44±1.05%, 91.89±0.70%, 90.28±0.95% and 87.00±0.75%, respectively. The maximum sensitivity of s1, s2 and s3 are all in level 2 with the value of 91.83±3.09%, 92.67±2.38% and 91.17±2.61%.

Table 7. Results of different decomposition levels.

| Level | S1 | S2 | S3 | OA |
|---|---|---|---|---|
| 1 | 88.83± 3.85 | 89.33± 2.25 | 90.17± 2.54 | 89.44± 1.05 |
| 2 | 91.83± 3.09 | 92.67± 2.38 | 91.17± 2.61 | 91.89± 0.70 |
| 3 | 90.33± 3.31 | 91.00± 2.85 | 89.50± 2.84 | 90.28± 0.95 |
| 4 | 87.00± 3.58 | 87.67± 2.63 | 86.33± 2.05 | 87.00± 0.75 |

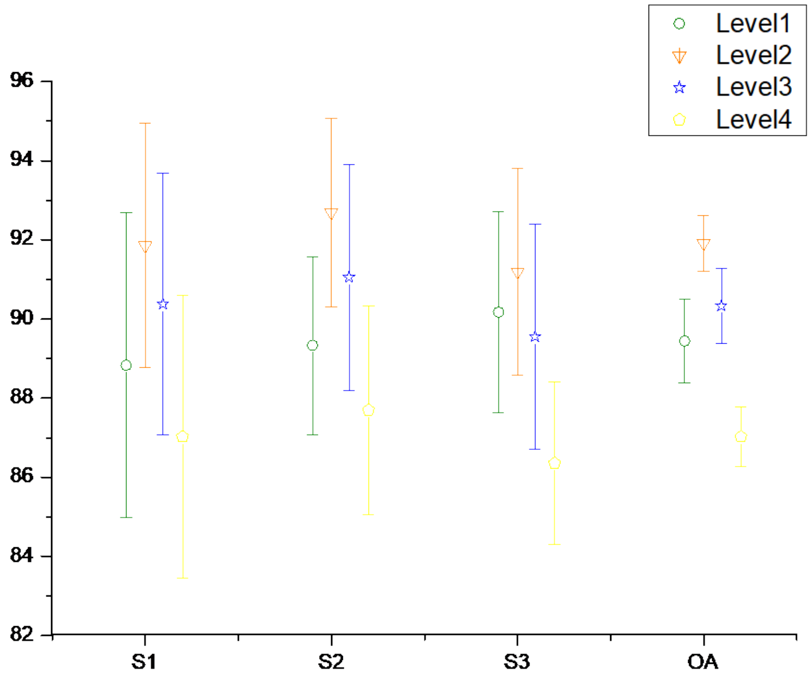


**Figure 10.** Error bar of various decomposition levels.

In Figure 10, we can observe that the data of sensitivity and overall accuracy peaked at 2-level SWE and gradually declined. Therefore, we chose the 2-level decomposition in our experiment. It is not hard to understand that 2-level SWE offers more information with seven sub-bands which are the same sizes of original brain image.

### 5.5. Comparison to State-of-the-art Approaches

We compared the model accuracy of this experiment with other current advanced methods: WE-GA [9], HMI [36] and SVM-PSO [10]. The comparison results are shown in Table 8. It shows HMI yielded an OA of 90.22±0.95%, the WE-GA yielded an OA of 77.47±1.17%, the SVM-PSO yielded an OA of 81.11±1.34% and the SWE-BBO yielded an OA of 86.17±0.41%. It is obvious from the numerical values that our method has a good performance.

Table 8. Comparison with other methods.

| Approach | Overall accuracy |
|---|---|
| HMI | 77.47±1.17 |
| WE-GA | 81.11±1.34 |
| SVM-PSO | 85.60±0.84 |
| **SWE-BBO** | **91.89±0.70** |

(Bold one is the best one)

## 6. Conclusions

In this paper, our team proposed to use SWE and single-hidden-layer neural network to constitute a system for detecting hearing loss, and to optimize it by BBO biogeography algorithm. The overall accuracy of HC, LHL and RHL is 91.89±0.70%. At the same time, we also analyzed the traditional model, elaborated our choice reason, and confirmed the optimal decomposition level 2, db4 and SWE through the accuracy. In the future we will explore more effective neural network searches to detect pathological images. In addition, we will actively seek more effective image preprocessing methods in order to achieve higher accuracy.

## References


[1] L. D. Landegger *et al.*, "Cytokine Levels in Inner Ear Fluid of Young and Aged Mice as Molecular Biomarkers of Noise-Induced Hearing Loss," (in English), *Front. Neurol.,* Article vol. 10, p. 14, Sep 2019, Art no. 977, doi: 10.3389/fneur.2019.00977.

[2] J. Yuebo, S. Xian, and T. Yan, "Efficacy and safety of acupuncture therapy for nerve deafness: a meta-analysis of randomized controlled trials," *International journal of clinical and experimental medicine,* vol. 8, no. 2, 2015.

[3] S. Wang, Z. Dong, G. Ji, Y. Zhang, P. Phillips, and J. Yang, "Exponential Wavelet Iterative Shrinkage Thresholding Algorithm for compressed sensing magnetic resonance imaging," (in eng), *Information Sciences: An International Journal,* vol. 322, pp. 115-132, 2015.

[4] P. Chen, "Computer-aided detection of left and right sensorineural hearing loss by wavelet packet decomposition and least-square support vector machine," *Journal of the American Geriatrics Society,* vol. 64, no. S2, 2016, Art no. S350.

[5] J. Li, "Texture Analysis Method Based on Fractional Fourier Entropy and Fitness-scaling Adaptive Genetic Algorithm for Detecting Left-sided and Right-sided Sensorineural Hearing Loss," *Fundamenta Informaticae,* vol. 151, no. 1-4, pp. 505-521, 2017.

[6] J. Ramirez, "Unilateral sensorineural hearing loss identification based on double-density dual-tree complex wavelet transform and multinomial logistic regression," *Integrated Computer-Aided Engineering,* doi: 10.3233/ICA-190605.

[7] D. R. Nayak, "MJaya-ELM: A Jaya algorithm with mutation and extreme learning machine based approach for sensorineural hearing loss detection," *Applied Soft Computing,* vol. 83, p. 105626, 2019/10/01/ 2019.

[8] M. Yang, "Preliminary Study on Unilateral Sensorineural Hearing Loss Identification via Dual-Tree Complex Wavelet Transform and Multinomial Logistic Regression," in *Natural and Artificial Computation for Biomedicine and Neuroscience, Pt I*, vol. 10337, J. M. F. Vicente, J. R. A. Sanchez, F. D. Lopez, J. T. Moreo, and H. Adeli Eds., (Lecture Notes in Computer Science. Cham: Springer International Publishing Ag, 2017, pp. 289-297.

[9] O. Profant, A. Škoch, Z. Balogová, J. Tintěra, J. Hlinka, and J. Syka, "Diffusion tensor imaging and MR morphometry of the central auditory pathway and auditory cortex in aging," *Neuroscience,* vol. 260, 2014.

[10] F. Liu, A. Nayeem, and A. Pereira, "Hearing Loss Detection Based on Wavelet Entropy and Genetic Algorithm," 2017/11 2017: Atlantis Press, in 2017 International Conference on Applied Mathematics, Modeling and Simulation (AMMS 2017), pp. 49-53.

[11] F. Bao and K. Nakamura, "Hearing Loss via Wavelet Entropy and Particle Swarm Optimized Trained Support Vector Machine," *DEStech Transactions on Engineering and Technology Research*, 2019.

[12] D. Bansal, K. Khanna, R. Chhikara, R. K. Dua, and R. Malhotra, "Classification of Magnetic Resonance Images using Bag of Features for Detecting Dementia," *Procedia Computer Science,* vol. 167, 2020.

[13] N. Nha, V. An, C. Inchan, and W. Kyoung-Jae, "A stationary wavelet entropy-based clustering approach accurately predicts gene expression," *Journal of computational biology : a journal of computational molecular cell biology,* vol. 22, no. 3, 2015.

[14] S. Du, A. Liu, A. Atangana, S. Wang, and Z. Lu, "Application of stationary wavelet entropy in pathological brain detection," (in eng), *Multimedia tools and applications,* vol. 77, no. 3, pp. 3701-3714, 2018.

[15] D. U. Wutsqa, D. U. Wutsqa, and M. Marwah, "Median Filter Noise Reduction of Image and Backpropagation Neural Network Model for Cervical Cancer Classification," (in eng), *Journal of Physics: Conference Series,* vol. 855, no. 1, p. 012062 (10pp), 2017, doi: 10.1088/1742-6596/855/1/012062.

[16] A. B. Ji, J. H. Pang, and H. J. Qiu, "Support vector machine for classification based on fuzzy training data," (in English), *Expert Syst. Appl.,* Article vol. 37, no. 4, pp. 3495-3498, Apr 2010, doi: 10.1016/j.eswa.2009.10.038.

[17] G. Hu, C. Yin, M. Wan, Y. Zhang, and Y. Fang, "Recognition of diseased Pinus trees in UAV images using deep learning and AdaBoost classifier," *Biosystems Engineering,* vol. 194, 2020.

[18] X. Jiang, S. C. Satapathy, L. Yang, S.-H. Wang, and Y.-D. Zhang, "A Survey on Artificial Intelligence in Chinese Sign Language Recognition," *Arabian Journal for Science and Engineering,* 2020/07/22 2020, doi: 10.1007/s13369-020-04758-2.

[19] D.-M. Pu, D.-Q. Gao, T. Ruan, and Y.-B. Yuan, "A novel learning algorithm of single-hidden-layer feedforward neural

networks," *Neural Computing and Applications,* vol. 28, no. 1, 2017.

[20] J. Liu, X. Jin, F. Dong, L. He, and H. Liu, "Fading channel modelling using single-hidden layer feedforward neural networks," (in eng), *Multidimensional systems and signal processing,* vol. 28, no. 3, pp. 885-903, 2017.

[21] A. K. Sangaiah, "Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization," *Neural Computing and Applications,* vol. 32, pp. 665-680, 2020, doi: 10.1007/s00521-018-3924-0.

[22] Y. Chen, "Cerebral micro-bleeding identification based on a nine-layer convolutional neural network with stochastic pooling," *Concurrency and Computation: Practice and Experience,* vol. 31, no. 1, p. e5130, 2020, doi: 10.1002/cpe.5130.

[23] V. V. Govindaraj, "High performance multiple sclerosis classification by data augmentation and AlexNet transfer learning model," *Journal of Medical Imaging and Health Informatics,* vol. 9, no. 9, pp. 2012-2021, 2019.

[24] C. Tang, "Cerebral Micro-Bleeding Detection Based on Densely Connected Neural Network," (in English), *Frontiers in Neuroscience,* Original Research vol. 13, 2019-May-17 2019, Art no. 422, doi: 10.3389/fnins.2019.00422.

[25] S. Xie, "Alcoholism Identification Based on an AlexNet Transfer Learning Model," (in English), *Frontiers in Psychiatry,* Original Research vol. 10, 2019-April-11 2019, Art no. 205, doi: 10.3389/fpsyt.2019.00205.

[26] G. Zhao, "Polarimetric synthetic aperture radar image segmentation by convolutional neural network using graphical processing units," *Journal of Real-Time Image Processing,* vol. 15, no. 3, pp. 631-642, 2018, doi: 10.1007/s11554-017-0717-0.

[27] C. Huang, "Multiple Sclerosis Identification by 14-Layer Convolutional Neural Network With Batch Normalization, Dropout, and Stochastic Pooling," (in English), *Frontiers in Neuroscience,* Original Research vol. 12, 2018-November-08 2018, Art no. 818, doi: 10.3389/fnins.2018.00818.

[28] K. Muhammad, "Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation," *Multimedia Tools and Applications,* vol. 78, no. 3, pp. 3613-3632, 2019, doi: 10.1007/s11042-017-5243-3.

[29] C. Pan, "Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU," *Journal of Computational Science,* vol. 28, pp. 1-10, 2018/09/01/ 2018.

[30] D.-w. Gong, Y. Zhang, and C.-l. Qi, "Environmental/economic power dispatch using a hybrid multi-objective optimization algorithm," *International Journal of Electrical Power and Energy Systems,* vol. 32, no. 6, 2009.

[31] K. S. Lee and Z. W. Geem, "A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice," *Computer Methods in Applied Mechanics and Engineering,* vol. 194, no. 36, 2004.

[32] D. Simon, "Biogeography-Based Optimization," (in English), *IEEE Trans. Evol. Comput.,* Article vol. 12, no. 6, pp. 702-713, Dec 2008, doi: 10.1109/tevc.2008.919004.

[33] X. Zhang *et al.*, "Improved Laplacian Biogeography-Based Optimization Algorithm and Its Application to QAP," *Complexity,* vol. 2020, 2020.

[34] Q. Niu, L. Zhang, and K. Li, "A biogeography-based optimization algorithm with mutation strategies for model parameter estimation of solar and fuel cells," *Energy Conversion and Management,* vol. 86, 2014.

[35] Y.-J. Li, "Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization," *Multimedia Tools and Applications,* vol. 77, no. 9, pp. 10393-10417, 2018, doi: 10.1007/s11042-016-4222-4.

[36] T. Lijun, Q. Yixuan, and A. Pereira, "Hu Moment Invariant: A New Method for Hearing Loss Detection," 2018/03 2018: Atlantis Press, pp. 412-416.