A Multiple Sclerosis Recognition via Hu Moment Invariant and Artificial Neural Network Trained by Particle Swarm Optimization

Ji Han[1] and Shou-Ming Hou[1,2(✉)]

[1] Henan Polytechnic University, Jiaozuo 454000, Henan, China
HanJi@home.hpu.edu.cn, housm@163.com
[2] Hebi Automotive Engineering Professional College,
Hebi 454030, Henan, China

Abstract. Multiple sclerosis can damage the central nervous system, and current drugs are difficult to completely cure symptoms. The aim of this paper was to use deep learning methods to increase the detection rate of multiple sclerosis, thereby increasing the patient's chance of treatment. We presented a new method based on hu moment invariant and artificial neural network trained by particle swarm optimization. Our method was carried out over ten runs of ten-fold cross validation. The experimental results show that the optimization ability of particle swarm optimization algorithm is superior to the genetic algorithm, simulated annealing algorithm and immune genetic algorithm. At the same time, compared with the HWT+PCA+LR method and the WE-FNN-AGA method, our method performs better in the performance of the detection.

Keywords: Multiple sclerosis · Hu moment invariant · Feedforward neural network · Particle swarm optimization

1 Introduction

Multiple sclerosis (MS) is an immune-mediated disease characterized by inflammatory demyelinating lesions of the white matter of the central nervous system [1]. Multiple sclerosis involves the effects of genetic, environmental and viral factors [2]. Multiple sclerosis is more likely to occur in young people, especially young women. Treating patients as early as possible can help slow the progression of the disease. Therefore, early detection of multiple sclerosis will benefit the treatment of patients.

In early stages of MS, there are neurodegenerative behaviors such as axons and loss of neurons. These behaviors can be used to determine permanent limb accumulation and cognitive impairment [3]. The treatment of MS mainly includes acute phase treatment and remission treatment. At present, the treatment plan for the acute phase is treated with high-dose glucocorticoid shock therapy, and the treatment for remission is recommended to use disease modification therapy.

As the advantages of deep learning methods in medical detection become more prominent, more researchers tend to use deep learning methods to perform multiple

Y.-D. Zhang et al. (Eds.): ICMTEL 2020, LNICST 327, pp. 254–264, 2020.
https://doi.org/10.1007/978-3-030-51103-6_22

sclerosis detection tasks. Alshayeji, Al-Rousan (2018) [4] proposed a high-efficiency detection system for multiple sclerosis based on neural network. Lopez (2017) [5] proposed a hybrid method combining principal component (PCA), Haar wavelet transform (HWT), and logistic regression (LR). Han and Hou (2019) [6] proposed a wavelet entropy (WE) with feedforward neural network (FNN), whose weights and biases were optimized through a new swarm intelligence method—adaptive genetic algorithm (AGA).

In the study, mathematical and morphological operations were used for feature extraction, and multi-layer feedforward neural networks were used to identify multiple sclerosis tissues. Currently, deep learning and transfer learning are widely used in medical image analysis [7–19]. Nevertheless, deep learning is not suitable for this study, because our dataset is too small.

In order to further improve the precision of detection, we use feedforward neural network in combination with hu moment invariant and particle swarm optimization algorithm to detect multiple sclerosis. The rest of the organization structure of this paper is as follows. Section 2 shows the source of the dataset used. Section 3 introduces the methods used to detect multiple sclerosis. Section 4 analyzes and discusses the experimental results. Section 5 summarizes the research contributions of the paper.

2 Dataset

Within our paper, the dataset we used were from Ref [17] and eHealth laboratory [20]. Table 1 shows the demographic characteristics of two datasets. 681 healthy slice brain images from 26 healthy controls and 676 MS slice brain images from 38 patients with multiple sclerosis were selected. All brain slice image data is guaranteed and confirmed by a professional brain surgeon. Therefore, the dataset used in the experiment is accurate and reliable.

Table 1. Demographic characteristics of two datasets.

Subject	Provider	NS1	NS2	Gender (m/f)	Age
Multiple sclerosis	eHealth laboratory	38	676	17/21	34.1 ± 10.5
Healthy control	Dr. Pan et al.	26	681	12/14	33.5 ± 8.3

(NS1 = number of subjects; NS2 = number of slice)

3 Methodology

The implementation of our detection method is based on Hu moment invariant (HMI), feedforward neural network, and particle swarm optimization (PSO). Among them, HMI is used to extract the enhanced dataset, feedforward neural network is used for feature learning and classification, and PSO is used to optimize the trained model because of its intelligent optimization ability.

3.1 HMI

In this study, image moment is used as the shape descriptor. For a 2D brain image $I(a, b)$, the raw moment M of order (p, q) is defined as

$$M_{pq} = \Sigma_a \Sigma_b a^p b^q I(a, b) \tag{1}$$

where u, v = 0, 1, 2,…

The central moment increases the translation invariance based on raw moment [21], and the central moment N is defined as

$$N_{pq} = \Sigma_a \Sigma_b [a - E(a)]^p [b - E(b)]^q I(a, b) \tag{2}$$

$$E(a) = \frac{M_{10}}{M_{00}} \tag{3}$$

$$E(b) = \frac{M_{01}}{M_{00}} \tag{4}$$

The normalized central moment increases the scale invariant based on the central moment [22], and the normalized central moment φ is defined as:

$$\varphi_{pq} = \frac{N_{pq}}{N_{00}\left(\frac{p+q}{2} + 1\right)} \tag{5}$$

The hu moment increases rotation invariant based on normalized central moment [23]. Seven Hu moment invariants $[\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7]$ are defined as:

$$\omega_1 = \varphi_{02} + \varphi_{20} \tag{6}$$

$$\omega_2 = 4\varphi_{11}^2 + (\varphi_{20} - \varphi_{02})^2 \tag{7}$$

$$\omega_3 = (3\varphi_{21} - \varphi_{03})^2 + (\varphi_{30} - 3\varphi_{12})^2 \tag{8}$$

$$\omega_4 = (\varphi_{03} + \varphi_{21})^2 + (\varphi_{12} + \varphi_{30})^2 \tag{9}$$

$$\begin{aligned} \omega_5 = {} & \left[3(\varphi_{12} + \varphi_{30})^2 - (\varphi_{03} + \varphi_{21})^2\right](\varphi_{03} + \varphi_{21})(3\varphi_{21} - \varphi_{03}) \\ & + \left[(\varphi_{12} + \varphi_{30})^2 - 3(\varphi_{03} + \varphi_{21})^2\right](\varphi_{12} + \varphi_{30})(\varphi_{30} - 3\varphi_{12}) \end{aligned} \tag{10}$$

$$\begin{aligned} \omega_6 = {} & 4\varphi_{11}(\varphi_{03} + \varphi_{21})(\varphi_{12} + \varphi_{30}) \\ & + \left[(\varphi_{12} + \varphi_{30})^2 - (\varphi_{03} + \varphi_{21})^2\right](\varphi_{20} - \varphi_{02}) \end{aligned} \tag{11}$$

$$\begin{aligned} \omega_7 = {} & \left[(\varphi_{12} + \varphi_{30})^2 - 3(\varphi_{03} + \varphi_{21})^2\right](\varphi_{12} + \varphi_{30})(3\varphi_{21} - \varphi_{03}) \\ & - \left[3(\varphi_{12} + \varphi_{30})^2 - (\varphi_{03} + \varphi_{21})^2\right](\varphi_{03} + \varphi_{21})(\varphi_{30} - 3\varphi_{12}) \end{aligned} \tag{12}$$

Hu moment invariance (HMI) is successfully applied to object recognition, image detection, visual tracking, shape measurement, etc. due to its characteristics of translation, scale, and rotation invariants. [24]. In the future, we shall test other feature extractors [25–32].

3.2 Feedforward Neural Network

Feedforward neural networks (FNN) usually consist of an input layer, a hidden layer, and an output layer. The nodes of each adjacent layer are fully linked, and the size of the link weight reflects the closeness of the relationship between the two nodes. FNN can only propagate in the direction from input to output. The output of any neuron in the same layer will not affect other neurons, and there is a mapping relationship between the layers [33]. The structure of the general feedforward neural network is shown in Fig. 1.

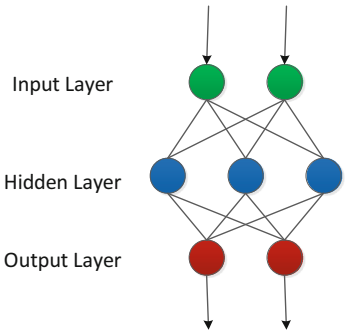

Fig. 1. Structure of feedforward neural network.

3.3 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is an algorithm for intelligent optimization by simulating the predation of birds [34]. Particles represent the solution to the problem, and fitness functions are used to evaluate the performance of the solution. During each iteration, local and global optimal solutions are selected to update the position of the particles. After several iterations, the particle will search for the optimal solution. The workflow of the PSO is shown in Fig. 2.

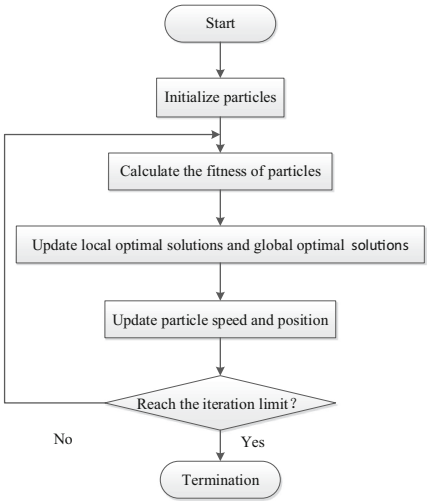

Fig. 2. Workflow of the PSO.

Assuming that there is only one optimal solution in region T, the position and velocity of each particle are updated according to Eqs. (13) and (14).

$$V_{it}^{n+1} = wV_{it}^{n} + a_1 r_1 (P_{best}^{n} - y_{it}^{n}) + a_2 r_2 (G_{best}^{n} - y_{it}^{n}) \tag{13}$$

$$Y_{it}^{n+1} = Y_{it}^{n} + V_{it}^{n+1} \tag{14}$$

where V_{it}^{n+1} represents the speed reached by the i-th particle in the $(n+1)$-th iteration, Y_{it}^{n} is the position of the i-th particle after nth iteration, w is inertia factor, a_1 and a_2 are acceleration constants, r_1 and r_2 are random numbers in the range [0, 1], t represents the position and speed of the t-th dimension. P_{best}, G_{best} represents the local optimal solution and the global optimal solution, respectively.

3.4 10-fold Cross Validation

The 10-fold cross validation in this paper is a method for assessing the accuracy of models built on datasets [35]. In 10-fold cross-validation, the data set is divided into 10 random subsets. In each iteration, one subset is selected as a test subset in turn, and the remaining nine subsets are used for training. After 10 iterations, the results of the verification are averaged to analyze the performance of the model [36]. This experiment will run 10 times 10-fold cross validation to ensure more accurate results.

4 Results and Discussions

4.1 Statistical Analysis

This proposed MS detection method "HMI-FNN-PSO" is implemented over 10 runs of 10-fold cross validation. Table 2 shows the results. Our method achieved satisfactory results on five indicators: sensitivity (Se), specificity (Sp), precision (Pr), accuracy (Ac), and F1 score (Fs). The test results of indicators Se, Sp, Pr, Ac, and Fs were $91.67 \pm 1.41\%$, $91.73 \pm 0.77\%$, $91.70 \pm 0.78\%$, $91.70 \pm 0.97\%$, and $91.67 \pm 1.00\%$.

Table 2. 10 runs of 10-fold cross validation of our method.

Run	Se	Sp	Pr	Ac	Fs
1	93.05	92.51	92.53	92.78	92.78
2	94.38	92.95	93.00	93.66	93.68
3	91.12	90.75	90.71	90.94	90.91
4	92.75	92.36	92.37	92.56	92.55
5	89.05	91.49	91.29	90.27	90.14
6	90.98	91.63	91.54	91.30	91.26
7	91.86	91.93	91.88	91.90	91.87
8	91.86	90.31	90.39	91.08	91.12
9	91.27	91.34	91.27	91.30	91.27
10	90.38	92.08	92.03	91.23	91.16
Average	91.67 ± 1.41	91.73 ± 0.77	91.70 ± 0.78	91.70 ± 0.97	91.67 ± 1.00

4.2 Training Algorithm Comparison

We compared PSO with three training algorithms: genetic algorithm (GA), simulated annealing (SA), and immune genetic algorithm (IGA). The results of 10 runs of 10-fold cross validation using GA, SA, and IGA are shown in Table 3, Table 4, and Table 5, respectively. Table 6 shows the results of the comparison and Fig. 3 shows the results in the form of a histogram. The experimental results confirm that the performance of our training algorithm is superior to the other three training algorithms, and our test standards are more than 90%.

Table 3. 10 runs of 10-fold cross validation using GA.

Run	Se	Sp	Pr	Ac	Fs
1	85.21	85.76	85.59	85.48	85.38
2	84.62	81.35	81.83	82.98	83.20
3	84.02	78.71	79.73	81.36	81.78
4	84.32	84.43	84.31	84.38	84.30
5	83.88	85.16	85.00	84.52	84.38
6	85.21	86.34	86.09	85.78	85.64
7	81.36	81.64	81.52	81.50	81.42
8	84.91	86.93	86.60	85.92	85.74
9	82.54	85.47	85.01	84.01	83.74
10	84.91	85.91	85.71	85.41	85.30
Average	84.10 ± 1.19	84.17 ± 2.54	84.14 ± 2.18	84.13 ± 1.60	84.09 ± 1.47

Table 4. 10 runs of 10-fold cross validation using SA.

Run	Se	Sp	Pr	Ac	Fs
1	79.44	78.71	78.75	79.07	79.09
2	81.07	79.14	79.51	80.10	80.25
3	82.10	82.96	82.72	82.53	82.41
4	77.22	81.20	80.31	79.22	78.73
5	81.66	82.38	82.17	82.02	81.90
6	78.25	76.07	76.47	77.16	77.33
7	81.07	78.71	79.07	79.88	80.06
8	77.37	80.47	79.73	78.92	78.53
9	79.59	77.68	77.97	78.63	78.77
10	80.47	81.94	81.56	81.21	81.00
Average	79.82 ± 1.66	79.93 ± 2.12	79.83 ± 1.83	79.87 ± 1.56	79.81 ± 1.53

Table 5. 10 runs of 10-fold cross validation using IGA.

Run	Se	Sp	Pr	Ac	Fs
1	87.57	88.40	88.23	87.99	87.90
2	84.32	85.76	85.44	85.04	84.88
3	87.43	86.49	86.53	86.96	86.97
4	88.61	90.16	89.98	89.39	89.28
5	86.54	86.93	86.79	86.74	86.66
6	85.06	83.26	83.46	84.16	84.25
7	83.58	81.06	81.41	82.31	82.48
8	84.17	84.58	84.42	84.38	84.30
9	86.83	85.61	85.70	86.22	86.25
10	86.69	89.43	89.09	88.06	87.87
Average	86.08 ± 1.60	86.17 ± 2.64	86.11 ± 2.48	86.12 ± 2.04	86.08 ± 1.97

Table 6. Comparison of different training algorithms.

Method	Se	Sp	Pr	Ac	Fs
GA	84.10 ± 1.19	84.17 ± 2.54	84.14 ± 2.18	84.13 ± 1.60	84.09 ± 1.47
SA	79.82 ± 1.66	79.93 ± 2.12	79.83 ± 1.83	79.87 ± 1.56	79.81 ± 1.53
IGA	86.08 ± 1.60	86.17 ± 2.64	86.11 ± 2.48	86.12 ± 2.04	86.08 ± 1.97
PSO (ours)	91.67 ± 1.41	91.73 ± 0.77	91.70 ± 0.78	91.70 ± 0.97	91.67 ± 1.00

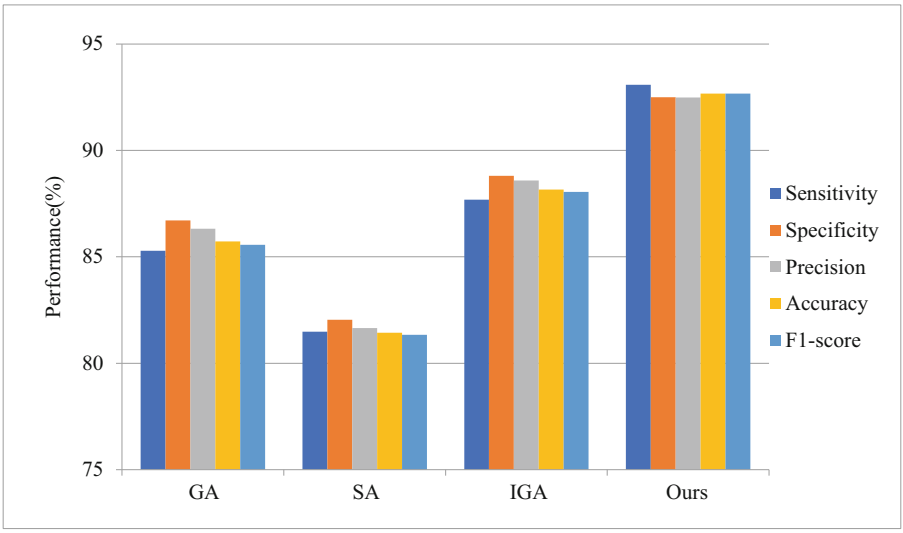

Fig. 3. Comparison of different training algorithms

4.3 Comparison with State-of-the-Art Algorithms

We compared our method with two state of the art approaches: one based on Haar wavelet transform (HWT), principal component analysis (PCA), and logistic regression (LR). Another based on wavelet entropy (WE), feedforward neural network (FNN), and adaptive genetic algorithm (AGA). Table 7 shows the results of the comparison. In Fig. 4, the comparison results are more intuitive in the form of a histogram. From an precision point of view, the experimental results of HWT+PCA+LR [5] and WE-FNN-AGA [6] were 87.51% and 86.32 ± 0.86%, respectively. However, the precision of our method is 91.70 ± 0.78%, so it is significantly better than the two state-of-the-art approaches.

Table 7. Comparison with state-of-the-art approaches.

Method	Se	Sp	Pr	Ac	Fs
HWT+PCA+LR [5]	89.04	90.24	87.51	89.72 ± 1.18	88.27
WE-FNN-AGA [6]	86.32 ± 1.09	86.36 ± 1.06	86.32 ± 0.86	86.34 ± 1.12	n/a
HMI-FNN-PSO (ours)	91.67 ± 1.41	91.73 ± 0.77	91.70 ± 0.78	91.70 ± 0.97	91.67 ± 1.00

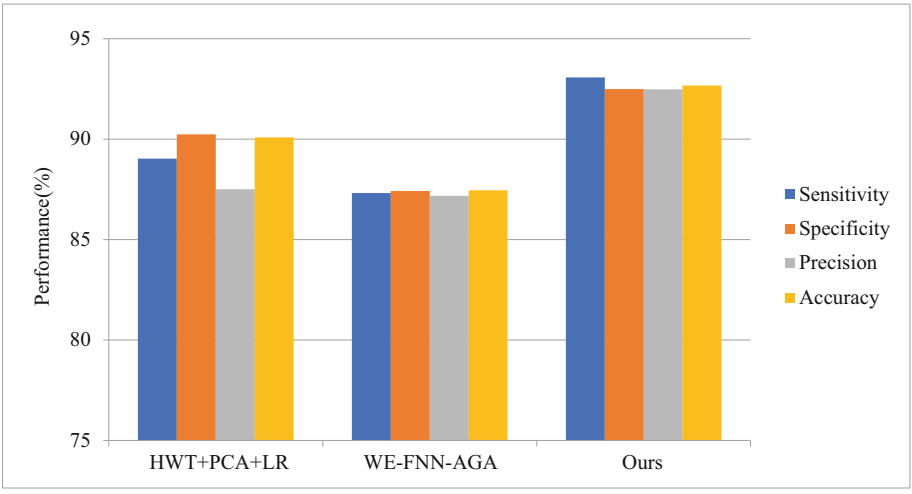

Fig. 4. Comparison to state-of-the-art approaches.

5 Conclusions

In this paper, we conduct research based on the detection of brain images of multiple sclerosis. Combine hu moment invariant and particle swarm optimization algorithms with feedforward neural networks. Our main research contributions are shown below:

(1) For the first time, we applied Hu Moment invariant and feedforward neural networks to the detection of multiple sclerosis, which improved the detection rate and helped patients with early treatment.
(2) In order to improve the optimization ability, a particle swarm algorithm is proposed and the performance is superior.
(3) Compared with other state-of-the-art approaches, the proposed method achieves the best performance in terms of index Se, Sp, Pr and Ac. Test results show that our detection performance is significantly better than the HWT+PCA+LR and WE-FNN-AGA approach.

In future research, we will further innovate in algorithms to improve detection accuracy and make more contributions to the early detection of multiple sclerosis.

Acknowledgement. This paper was supported by Key research and development and technology promotion projects in Henan Province, China (No. 172102210273, 182102210086, 182102310629), Youth backbone training program for colleges and universities in Henan, China (No. 2018GGJS298).

References

1. Barro, C., et al.: Serum neurofilament as a predictor of disease worsening and brain and spinal cord atrophy in multiple sclerosis. Brain **141**(8), 2382–2391 (2018)
2. Tintore, M., et al.: Treatment of multiple sclerosis—success from bench to bedside. Nat. Rev. Neurol. **15**(1), 53–58 (2019)
3. Pietroboni, A.M., et al.: The loss of macular ganglion cells begins from the early stages of disease and correlates with brain atrophy in multiple sclerosis patients. Mult. Scler. J. **25**(1), 31–38 (2019)
4. Alshayeji, M.H., et al.: An efficient multiple sclerosis segmentation and detection system using neural networks. Comput. Electr. Eng. **71**, 191–205 (2018)
5. Lopez, M.: Multiple sclerosis slice identification by haar wavelet transform and logistic regression. Adv. Eng. Res. **114**, 50–55 (2017)
6. Han, J., Hou, S.-M.: Multiple sclerosis detection via wavelet entropy and feedforward neural network trained by adaptive genetic algorithm. In: Rojas, I., Joya, G., Catala, A. (eds.) IWANN 2019. LNCS, vol. 11507, pp. 87–97. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-20518-8_8
7. Govindaraj, V.V.: High performance multiple sclerosis classification by data augmentation and AlexNet transfer learning model. J. Med. Imaging Health Inform. **9**(9), 2012–2021 (2019)
8. Jiang, X.: Chinese sign language fingerspelling recognition via six-layer convolutional neural network with leaky rectified linear units for therapy and rehabilitation. J. Med. Imaging Health Inform. **9**(9), 2031–2038 (2019)
9. Yu, X., Zeng, N., Liu, S., Zhang, Y.-D.: Utilization of DenseNet201 for diagnosis of breast abnormality. Mach. Vis. Appl. **30**(7–8), 1135–1144 (2019)
10. Li, Z.: Teeth category classification via seven-layer deep convolutional neural network with max pooling and global average pooling. Int. J. Imaging Syst. Technol. **29**, 577–583 (2019). https://doi.org/10.1002/ima.22337
11. Hong, J.: Detecting cerebral microbleeds with transfer learning. Mach. Vis. Appl. **30**(7–8), 1123–1133 (2019)
12. Tang, C.: Cerebral micro-bleeding detection based on densely connected neural network. Front. Neurosci. **13**, 422 (2019)
13. Xie, S.: Alcoholism identification based on an AlexNet transfer learning model. Front. Psychiatr. **10**, 205 (2019)
14. Jia, W.: Five-category classification of pathological brain images based on deep stacked sparse autoencoder. Multimed. Tools Appl. **78**(4), 4045–4064 (2017)
15. Muhammad, K.: Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. Multimed. Tools Appl. **78**(3), 3613–3632 (2019)
16. Sangaiah, A.K.: Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization. Neural Comput. Appl. **32**(3), 665–680 (2019). https://doi.org/10.1007/s00521-018-3924-0
17. Pan, C.: Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU. J. Comput. Sci. **28**, 1–10 (2018)

18. Hou, X.-X.: Seven-layer deep neural network based on sparse autoencoder for voxelwise detection of cerebral microbleed. Multimed. Tools Appl. **77**(9), 10521–10538 (2018)
19. Pan, C.: Abnormal breast identification by nine-layer convolutional neural network with parametric rectified linear unit and rank-based stochastic pooling. J. Comput. Sci. **27**, 57–68 (2018)
20. MRI Lesion Segmentation in Multiple Sclerosis Database, in eHealth Laboratory, University of Cyprus. http://www.medinfo.cs.ucy.ac.cy/index.php/downloads/datasets
21. Wang, Y.Q., et al.: The optimal fractional S transform of seismic signal based on the normalized second-order central moment. J. Appl. Geophys. **129**, 8–16 (2016)
22. Yang, J.: Pathological brain detection in MRI scanning via Hu moment invariants and machine learning. J. Exp. Theor. Artif. Intell. **29**(2), 299–312 (2017)
23. Hu, M.-K.: Visual pattern recognition by moment invariants. IRE Trans. Inf. Theory **8**(2), 179–187 (1962)
24. Hou, X.-X.: Alcoholism detection by medical robots based on Hu moment invariants and predator-prey adaptive-inertia chaotic particle swarm optimization. Comput. Electr. Eng. **63**, 126–138 (2017)
25. Ramirez, J.: Unilateral sensorineural hearing loss identification based on double-density dual-tree complex wavelet transform and multinomial logistic regression. Integr. Comput.-Aided Eng. **26**, 411–426 (2019). https://doi.org/10.3233/ICA-190605
26. Yang, J.: An adaptive encoding learning for artificial bee colony algorithms. J. Comput. Sci. **30**, 11–27 (2019)
27. Gorriz, J.M.: Multivariate approach for Alzheimer's disease detection using stationary wavelet entropy and predator-prey particle swarm optimization. J. Alzheimer's Dis. **65**(3), 855–869 (2018)
28. Sun, J.: Preliminary study on angiosperm genus classification by weight decay and combination of most abundant color index with fractional Fourier entropy. Multimed. Tools Appl. **77**(17), 22671–22688 (2018)
29. Zhao, G.: Smart pathological brain detection by synthetic minority oversampling technique, extreme learning machine, and Jaya algorithm. Multimed. Tools Appl. **77**(17), 22629–22648 (2018)
30. Lu, S.: Pathological brain detection in magnetic resonance imaging using combined features and improved extreme learning machines. J. Med. Imaging Health Informat. **8**, 1486–1490 (2018)
31. Li, Y.-J.: Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. Multimed. Tools Appl. **77**(9), 10393–10417 (2018)
32. Kong, F.Q.: Ridge-based curvilinear structure detection for identifying road in remote sensing image and backbone in neuron dendrite image. Multimed. Tools Appl. **77**(17), 22857–22873 (2018)
33. Chowdhury, N.: A comparative analysis of feed-forward neural network & recurrent neural network to detect intrusion. In: 2008 International Conference on Electrical and Computer Engineering, pp. 488–492. IEEE (2008)
34. Eberhart, R., et al.: A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, MHS 1995, pp. 39–43 (1995)
35. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: IJCAI, Montreal, Canada, pp. 1137–1145 (1995)
36. Sharma, M., et al.: A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension. Pattern Recogn. Lett. **94**, 172–179 (2017)