Adaptive Neuro-Fuzzy Inference System-Based Chaotic Swarm Intelligence Hybrid Model for Recognition of Mild Cognitive Impairment from Resting-State fMRI

Ahmed M. Anter[1,2](✉) and Zhiguo Zhang[1]

[1] School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen 518060, China
anter@szu.edu.cn

[2] Faculty of Computers and Information, Beni-Suef University, Benisuef 62511, Egypt
ahmed_anter@fcis.bsu.edu.eg

Abstract. Individuals diagnosed with mild cognitive impairment (MCI) are at high risk of transition to Alzheimer's disease (AD). Resting-state functional magnetic resonance imaging (rs-fMRI) is a promising neuroimaging technique for identifying patients with MCI. In this paper, a new hybrid model based on Chaotic Binary Grey Wolf Optimization Algorithm (CBGWO) and Adaptive Neuro-fuzzy Inference System (ANFIS) is proposed; namely (CBGWO-ANFIS) to diagnose the MCI. The proposed model is applied on real dataset recorded by ourselves and the process of diagnosis is comprised of five main phases. Firstly, the fMRI data are preprocessed by sequence of steps to enhance data quality. Secondly, features are extracted by localizing 160 regions of interests (ROIs) from the whole-brain by overlapping the Dosenbach mask, and then fractional amplitude of low-frequency fluctuation (fALFF) of the signals inside ROIs is estimated and used to represent local features. Thirdly, feature selection based non-linear GWO, chaotic map and naive Bayes (NB) are used to determine the significant ROIs. The NB criterion is used as a part of the kernel function in the GWO. CBGWO attempts to reduce the whole feature set without loss of significant information to the prediction process. Chebyshev map is used to estimate and tune GWO parameters. Fourthly, an ANFIS method is utilized to diagnose MCI. Fifthly, the performance is evaluated using different statistical measures and compared with different met-heuristic algorithms. The overall results indicate that the proposed model shows better performance, lower error, higher convergence speed and shorter execution time with accuracy reached to 86%.

Keywords: rs-fMRI · MCI · Optimization · Swarm intelligence · ANFS · Chaos theory

I. Rekik et al. (Eds.): PRIME 2019, LNCS 11843, pp. 23–33, 2019.
https://doi.org/10.1007/978-3-030-32281-6_3

1 Introduction

Alzheimer's disease (AD) is the most common cause of dementia and it is a heavy burden to the patients and the society. Individuals with MCI are considered to be at a much higher risk of developing AD, with a yearly conversion rate of 15–20%. Therefore, early detection of MCI is important for possible delay of the progression of MCI to moderate and severe stages. Recently, using rs-fMRI for detection of MCI has gained popularity because of the simple experimental procedure and readily available data analytics tools [1]. Particularly, many studies have investigated neuroimaging and machine learning (ML) techniques for the early detection of MCI and conversion from MCI to AD on the benchmark datasets [2–6].

Generally, ML has been successful implemented in many application areas. Despite the efficacy of ML on the diagnosis of AD and MCI, there are still various problems, such as falling in local optimum, not being able to converge to optimum solution, sensitivity to noise, and data uncertainty. Among various difficulties in detection of MCI from rs-fMRI, the most challenging one should be the curse of dimensionality [7]. Discovering valuable hidden knowledge in large scale data has become a pressing matter. Generally, there are two famous approaches for feature selection: filter approach is based on criterion to score features and wrapper approach contains ML techniques as a part of the evaluation process. These methods suffer from recession and stuck in local minima as well as computationally expensive. In order to further improve the effect of feature selection, the global optimization algorithm is needed [8].

In this paper, a new hybrid intelligent optimization model based on binary version of grey wolf optimizer (GWO), Chebyshev chaotic map, and ANFIS is proposed for rs-fMRI-based MCI diagnosis. The proposed model comprised of five main phases. (1) Preprocessing: fMRI data are sliced, aligned, filtered, normalized, and smoothed to enhance data quality. (2) Feature extraction: 160 regions of interest (ROIs) are localized from the whole-brain fMRI data by overlapping the Dosenbach mask. Afterwards, fractional amplitude of low-frequency fluctuation (fALFF) of the signals inside ROIs are estimated and used to represent local features. (3) Feature selection strategy: non-linear grey wolf optimizer (GWO) based on Chebyshev chaotic maps and naive Bayes (CBGWO) is used to determine the significant ROIs based power spectrum features. GWO is a metaheuristic optimization technique inspired by the hunting strategy and behavior of the grey wolves. The integration between these techniques is able to improve the classification accuracy of the MCI dementia. (4) MCI classification: hybrid fuzzy logic and artificial neural network (ANN) called (ANFIS) method is utilized to diagnose MCI based on the features selected from CBGWO. (5) Performance evaluation: different evaluation criteria are used for feature selection and classification phases. This model utilizes the strong ability of the global optimization of the GWO which avoids the sensitivity of local optimization. In order to minimize error, maximize classification accuracy, reduce erroneous clinical diagnostic. The following research problems for features selection and optimization will be addressed in this study.

– How to maintain several agents involved in the architecture?
– How to plan an integrated solution for a given MCI problem?
– How to avoid over-fitting using the integrated optimized model?
– How can get best local optima in the model without stuck in local minima?
– How to converge fast and decreasing time consuming?
– Is the global optimization algorithms need to parameters tuning?

The remaining of this paper is organized as follows. Section 2 presents the background knowledge and the newly-developed CBGWO model. Section 3 introduces the fMRI dataset and the pipeline of CBGWO-ANFIS model for MCI diagnosis. Experimental results are presented in Sect. 4. Finally, Sect. 5 concludes this paper and presents the future work.

2 Methodology

2.1 Grey Wolf Optimization Algorithm (GWO)

GWO is a meta-heuristic technique, which was presented in 2014 [9]. GWO is inspired from the hunting task and behavior of the grey wolves. Grey wolves are structured in a group from five to eleven members. They regulate the set into four kinds of wolves' categories (alpha (α), beta (β), omega (ω), delta (δ)) through hunting process, to remain the discipline in a group. The mathematical model that represents the strategy process of encircling the prey through the hunting is expressed with details in [9]:

$$X_{t+1} = X_{p,t} - A * D \tag{1}$$

$$d = |CX_{p,t} - X_t| \tag{2}$$

$$A = 2br_1 - b \tag{3}$$

$$C = 2r_2 \tag{4}$$

where X_{t+1} is the position of the wolf, $X_{p,t}$ is the position of the prey, d is a variation vector, r_1 and r_2 are random parameters between [0 1]. A and C are the main two coefficient vectors used to control the exploration and exploitation of GWO, and b is a linearly reducing vector from 2 to 0 over course of iterations, indicated in Eq. (5).

$$b = 2 - 2\left(\frac{\boldsymbol{t}}{\boldsymbol{T}_{max}}\right) \tag{5}$$

where t is the current iteration, and T_{max} is the maximum number of iterations.

2.2 The New Hybrid Kernel-Based CBGWO Model

The main challenging task in GWO is how to balance between exploration and exploitation with random distribution behavior. For nonlinear features, exploration and exploitation are in conflict and there is no clear boundary between them. For this reason, GWO suffers from low convergence and falling in local minima. Therefore, Chebyshev chaotic map (CCM) [13] is used for updating GWO position instead of the random behavior. In the new CBGWO method, CCM and naïve Bayes (NB) are used to tackle the problems of GWO.

GWO Dynamic Parameters Tuning: The main problem in GWO algorithm is a random distribution in dimensional space, which leads to the unbalance between the two main factors exploration and exploitation due to the stochastic nature of GWO and it can fall into the local minima. In this paper CCM is used to improve the performance of CBGWO in terms of convergence speed and local minima avoidance. In the standard GWO, two essential parameters *(A, C)* are used to control the behavior of the algorithm. CCM is utilized to tune these parameters by replacing the random generators variables as shown in Eqs. (6–8).

$$b_{\mathrm{CCM}} = \{(a_i - t)\left(\frac{a_f - a_i}{Max_{iter}}\right).CCM\,(t) \tag{6}$$

$$A_{\mathrm{ChoasVec}} = 2b.CCM\,(t) - b \tag{7}$$

$$C_{\mathrm{ChoasVec}} = 2.CCM(t) \tag{8}$$

where a_i, a_f are the elementary and latest values adjusted as 2, 0 respectively, t is the current iteration and Max_{iter} are the maximal number of iterations. b is linearly decreased from 2 to 0 based on CCM. CCM(t) is Chebyshev choatic map which is non-invertible, non-repetition and ergodicity of chaos. CCM is sequence of one dimensional choatic vector at iteration t and can be expressed by the following equation.

$$X_{i+1} = cos\left(tcos^{-1}\,(w_i)\,\right) \tag{9}$$

where i is an iteration number and the chaotic time series $\mathrm{x}_i \in [0, 1]$ are obtained.

NB-Based Kernel/Fitness Function: To select the discriminative features from the search process, non-parametric kernel density estimation is utilized as a learning tool using naïve Bayes for evaluating the predictive potential of the features from four different perspectives; size of the feature subset, classification accuracy, stability, and generalization of the feature set [10].

The following objective/fitness function is used to calculate the fitness value for each wolf in the search space.

$$Fitness = (1 - \omega) * (1 - Acc) + \omega * \frac{As}{(Al - 1)} \tag{10}$$

where ω is an initialized factor equal to 0.01, Acc is the accuracy of corrected data classified to the predicted label obtained from NB classifier, As is the summation of candidate attributes, and Al is the attribute length. The goodness of each wolf position in the feature space is assessed by this objective function. The main criteria are satisfied by minimizing the objective function.

3 Application on fMRI-Based MCI Recognition

3.1 Participants and Data Acquisition

All Individuals were diagnosed and scanned by expert doctors with more than 20 years' experience in dementia and radiology. The fMRI dataset was from the First Affiliated Hospital of Guangxi University of Chinese Medicine using Simens Magnetom Verio 3.0T°. The repetition time (TR) = 2000 ms and echo time (TE) = 30 ms, field of view (FOV) = 240 mm × 240 mm, imaging matrix = 64 × 64 slice, slice thickness = 5 mm, slices = 31, flip angle = 90°, total volumes = 180. In this study a total 127 subjects, 62 MCI (18 females, age: 66.66 ± 7.44; 44 males, age: 63.73 ± 5.94) and 65 NC (24 females, age: 66 ± 4.99; 41 males, age 64.12 ± 6.059), were recruited. All patients have written informed consent and informed about experimental content.

3.2 Pipeline and Phases of CBGWO-ANFIS

The aims of this paper are to develop new hybrid intelligence optimization model for MCI diagnosis from resting-state fMRI data. Figure 1 shows the schematic pipeline of the proposed CBGWO-ANFIS optimization model for MCI diagnosis. Basically, it is constructed using five main phases as follow:

1. **Pre-processing phase.** The fMRI preprocessing procedures include: removal of the first 5 volumes, slice timing, and head motion correction. The data are sliced, realigned, normalized, and smoothed to reduce the effects of bad normalization and to increase the signal-to-noise-ratio. Default mask (brainmask.nii) was used with a threshold of 50%. Several nuisance signals are regressed out from the data by utilizing the Friston 24-parameters model. Also, other resources of spurious variance are removed from the data through linear regression.
2. **Feature extraction phase.** 160 ROIs are localized using the Dosenbach pattern. fMRI signals inside the ROIs are presented as time-series. The fractional-ALFF (f-ALFF) method is used to transform time-series data into frequency domain and the power spectrum was obtained. The band-pass filter is performed to reduce the influences of low-frequency drift and high-frequency physiological noise. Then, principal component analysis (PCA) method is used to decrease the high correlation of this non-linear data.
3. **Feature selection phase.** CBGWO based on naive Bayes is used to determine the significant attributes in order to obtain the optimal features that represent the best ROIs for MCI. CCM is used to estimate and tune GWO parameters which dominate the exploration and exploitation rate.

4. **Classification phase.** The ANFIS method is used to provide an effective learning solution for MCI diagnosis [11].
5. **Performance evaluation phase.** The performance of the proposed model is evaluated using different statistical measures for feature selection and classification: Mean fitness (μ_f), Best fitness (BF), Standard deviation (Std), and Average Attributes Selected (AAS) for feature selection process. Accuracy (ACC), Precision, Recall, F1-score, Informedness (*Info*), Markedness (*Mark*), Matthews Correlation Coefficient (MCC), and kappa for classification process [12].

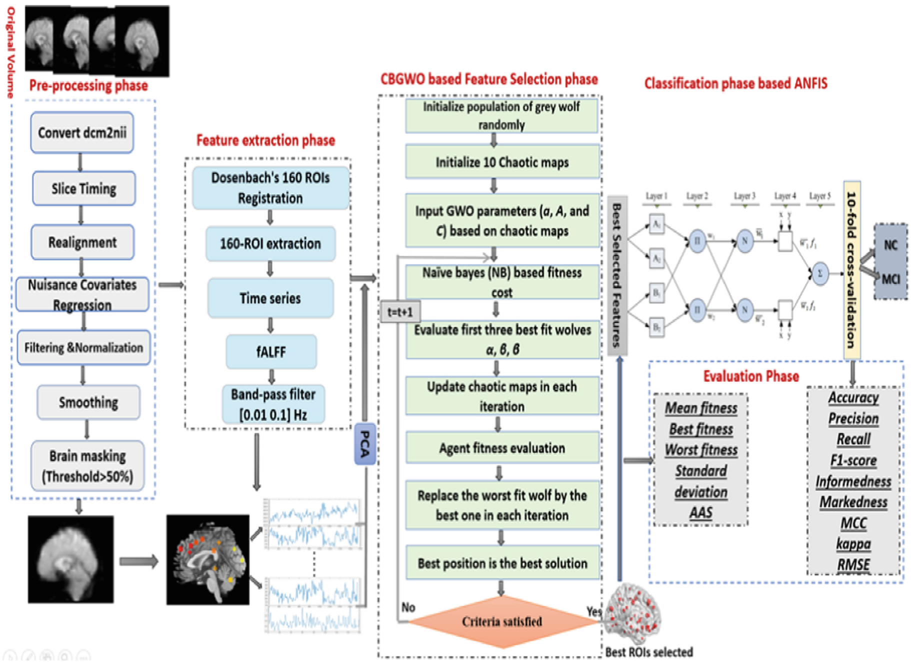

Fig. 1. General pipeline for the proposed model.

4 Experimental Results and Discussion

The proposed model was developed and tested via MATLAB (R2018b, The MathWorks, Inc.) on an Intel®CoreTM I5-6500 CPU 3.2 GHz processor and 16 GB RAM using Windows10 64-bit.

4.1 CBGWO-Based Feature Selection

Table 1 shows the best (BF) score, mean (μ), standard derivation (Std), number of selected features (SF), accuracy (ACC), and time consuming (CPU average time) obtained using BGWO based on CCM. We can see that, the optimal results (with the highest accuracy) are achieved using the following paramters setting; number of Agents (5), Max number of iterations (10), and Lower and Upper limits for binary search space [0,1].

The evidence of the convergence speed is provided in Fig. 2. The CBGWO algorithm has a strong robustness and has faster convergence speed to find the best solutions in less than 10 iterations. It is important to know how the kernel function decreases along the iterations and assist GWO to avoid local minima in the complex fMRI dataset.

Table 1. Results of the proposed BGWO model-based on NB.

Chaotic	*BF*	μ	*Std*	*SF*	*AAS*	*Time/s*	*ACC*
CCM	0.0371	0.0592	0.0153	82	0.513	31.000	0.8596

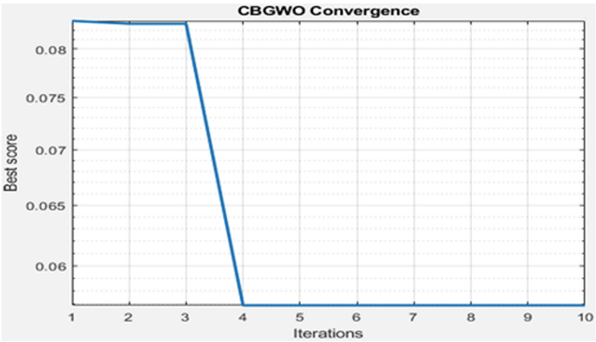

Fig. 2. BGWO algorithm convergence curve using CCB.

4.2 Diagnosis Based on ANFIS

Table 2 shows the accuracy results using the proposed model based on ANFIS and CCM with 10-fold cross validation (CV). We can see that, the highest accuracy achieved in almost all validation partitions with accuracy average ≈ 0.86. Moreover, Table 3 shos different statistical measures that used to validate the proposed model for MCI diagnosis. The accuracy achieved by Matthews correlation coefficient *MCC* ≈ 0.713, *F1-score* ≈ 0.84 which represents a perfect precision and recall even if the data is imbalanced, k is 0.69 which represent good agreement based on Kappa coefficient criteria $0.60 < k > 0.80$, *Info* ≈ 0.71 which measures of how informed the model about positives and negatives, and *Mark* ≈ 0.73 which measures trust worthiness of positive and negative predictions by the model.

Table 2. Accuracy results of the proposed CBGWO-ANFIS using 10-fold CV.

Ch.	CV_1	CV_2	CV_3	CV_4	CV_5	CV_6	CV_7	CV_8	CV_9	CV_{10}	Aver
CCM	0.8462	0.9167	1	0.7692	0.9231	1	0.6154	0.8333	1	0.6923	0.8596

Table 3. Results of CBGWO-ANFIS using different statistical measures.

Map	Acc	Recall	Precis.	F1-score	MCC	Kappa	Inform.	Mark.
CCM	0.8596	0.9257	0.7856	0.8395	0.7129	0.6900	0.7065	0.7273

In comparison with recent meta-heuristics optimization algorithms based on the performance of the selected features from rs-fMRI based MCI dataset. These algorithms are Whale optimization Algorithm (WOA) [13], Ant Lion Optimization (ALO) [15], Crow search optimization algorithm (CSA) [16], BGWO [17] algorithm, and BAT [18] algorithm. From Fig. 3, it is clearly seen that the CBGWO improved the accuracy for MCI prediction in comparison with WOA, ALO, BAT, CSA, and GWO algorithms with its chaotic maps. Further, since the BGWO-ANFIS is slightly better than CSA+Circle, we also performed 10 independent runs and afterward we used paired-sample t-test to compare these two algorithms. Results showed that CBGWO-ANFIS is significantly higher than CSA+Circle in the accuracy $(p = 0.03)$. From the results, we can see that, the proposed model achieved high prediction and quicker in locating the optimal solution. In general, it can search for the optimal solution in less than ten iterations.

Moreover, the proposed model compared with different machine learning (ML) techniques such as naïve Bayes (NB), Support vector Machine (SVM) with (Quadratic and Cubic), ANFIS, Logistic Regression (LR), Linear Discriminate Analysis (LDA), K-Nearest Neighbor (KNN), and Ensemble bagged Trees (EBT). All techniques are performed on the same MCI rs-fMRI dataset using 10-fold cross validation for training and testing sets. The results of these comparisons are shown in Fig. 4. It can be seen that, the more accurate results are achieved by using the proposed algorithm with high prediction accuracy ACC ($\approx$0.86), Sensitivity ($\approx$0.93), Specificity ($\approx$0.78), Precision ($\approx$0.79), and False positive rate (FPR) ($\approx$0.22), followed by the results achieved by ANFIS with ACC ($\approx$0.68), and followed by the results achieved by weighted KNN with ACC ($\approx$0.66). Further details about other methods can be seen in Fig. 4. In conclusion, the proposed model based ANFIS superior to conventional ML algorithms.

To validate the robustness of the proposed CBGWO-ANFIS model on other rs-fMRI dataset, we further examined the performance of the proposed model on a another dataset from ADNI [14]. Figure 5 shows the comparisons between CBGWO-ANFIS and several ML methods when being used ADNI dataset for classification of MCI patients from NC. As we can see, the proposed approach

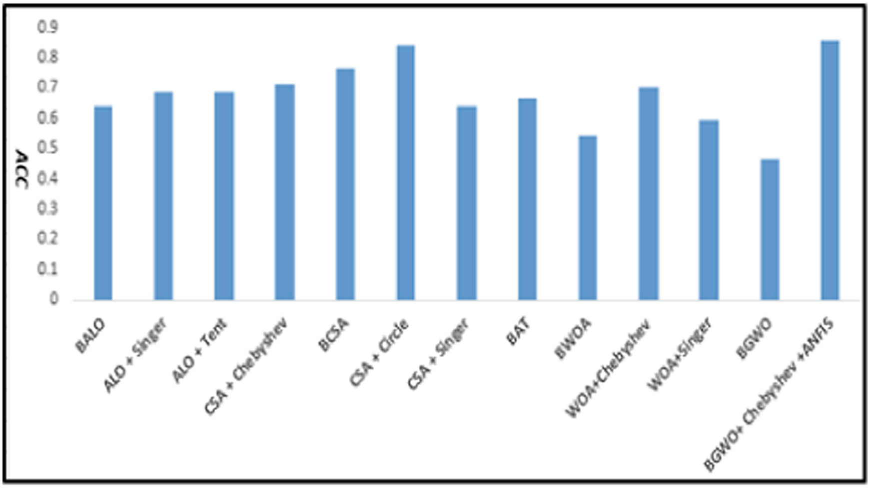

Fig. 3. Graphical representation accuracy results for different meta-heuristic algorithms.

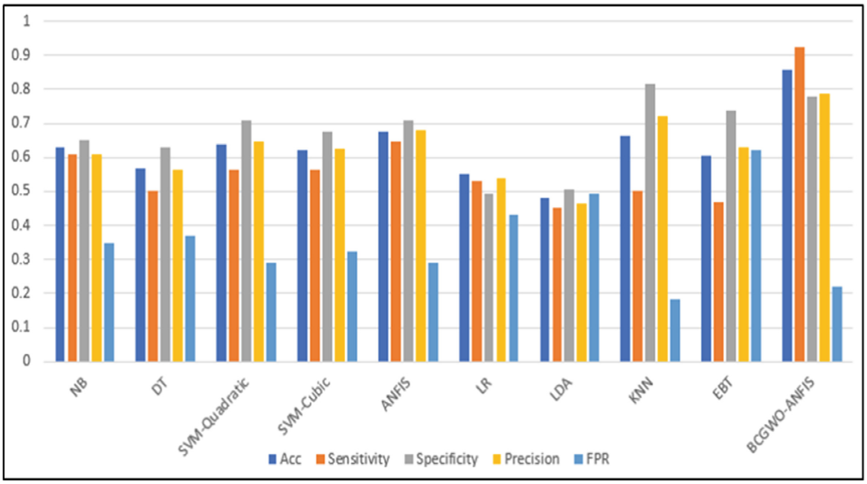

Fig. 4. Accuracy results for different ML techniques using different measures.

achieves the highest accuracy of 0.8232 and the highest sensitivity of 1. The specificity, precision and FPR of the proposed approach are also close to the best ones. These results further validate the better performance of the proposed CBGWO-ANFIS approach.

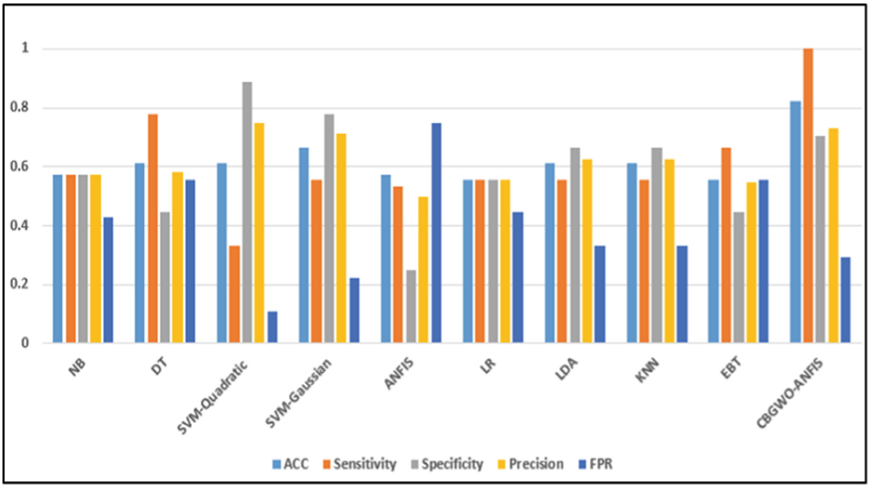

Fig. 5. Performance results and comparisons between CBGWO-ANFIS and other ML methods on ADNI dataset.

5 Conclusion and Future Work

In this paper, we proposed a novel MCI diagnosis model based on hybrid binary grey wolf optimizer, chaotic map and ANFIS algorithm (CBGWO-ANFIS). The experimental results demonstrate that the proposed hybrid model is reliable for MCI prediction especially with regard to high-dimensional data pattern recognition. The performance of the proposed model was measured on 127 subjects from real dataset. The experimental results show that the highest results obtained by the proposed model using 30 epochs and 10-fold cross validation. In a direct comparison, the proposed model exhibited better performance than bio-inspired optimization algorithms in literature. In future, the proposed model will be tested on more types of data (such as multimodal MRI), data from other centers, and even data of other brain diseases.

References

1. Vieira, S., Pinaya, W.H., Mechelli, A.: Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications. Neurosci. Biobehav. Rev. **74**, 58–75 (2017)
2. Suk, H.I., Wee, C.Y., Lee, S.W., Shen, D.: State-space model with deep learning for functional dynamics estimation in resting-state fMRI. NeuroImage **129**, 292–307 (2016)
3. Hosseini-Asl, E., Keynton, R., El-Baz, A.: Alzheimer's disease diagnostics by adaptation of 3D convolutional network. In 2016 IEEE International Conference on Image Processing (ICIP), pp. 126–130, September 2016

4. Han, X., Zhong, Y., He, L., Yu, P.S., Zhang, L.: The unsupervised hierarchical convolutional sparse auto-encoder for neuroimaging data classification. In: Guo, Y., Friston, K., Aldo, F., Hill, S., Peng, H. (eds.) BIH 2015. LNCS (LNAI), vol. 9250, pp. 156–166. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-23344-4_16
5. Kim, J., Calhoun, V.D., Shim, E., Lee, J.H.: Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. Neuroimage **124**, 127–146 (2016)
6. Grothe, M., Heinsen, H., Teipel, S.: Longitudinal measures of cholinergic forebrain atrophy in the transition from healthy aging to Alzheimer's disease. Neurobiol. Aging **34**(4), 1210–1220 (2013)
7. Li, Y., et al.: Abnormal resting-state functional connectivity strength in mild cognitive impairment and its conversion to Alzheimer's disease. Neural Plast. **2016**, 4680972 (2016). https://doi.org/10.1155/2016/4680972
8. Faris, H., Aljarah, I., Al-Betar, M.A., Mirjalili, S.: Grey wolf optimizer: a review of recent variants and applications. Neural Comput. Appl. **30**(2), 413–435 (2018)
9. Mirjalili, S., Mirjalili, S.M., Lewis, A.: Grey wolf optimizer. Adv. Eng. Softw. **69**, 46–61 (2014)
10. Li, T., Li, J., Liu, Z., Li, P., Jia, C.: Differentially private naive bayes learning over multiple data sources. Inf. Sci. **444**, 89–104 (2018)
11. Hasanipanah, M., Amnieh, H.B., Arab, H., Zamzam, M.S.: Feasibility of PSO-ANFIS model to estimate rock fragmentation produced by mine blasting. Neural Comput. Appl. **30**(4), 1015–1024 (2018)
12. Alam, S., Dobbie, G., Koh, Y.S., Riddle, P., Rehman, S.U.: Research on particle swarm optimization based clustering: a systematic review of literature and techniques. Swarm Evol. Comput. **17**, 1–13 (2014)
13. Anter, A., Gupta, D., Castillo, O.: novel parameter estimation in dynamic model via fuzzy swarm intelligence and chaos theory for faults in wastewater treatment plant. Soft Comput. 1–19 (2019). https://doi.org/10.1007/s00500-019-04225-7
14. ADNI. http://adni.loni.usc.edu/. Accessed April 2019
15. Mafarja, M., Eleyan, D., Abdullah, S., Mirjalili, S.: S-shaped vs. V-shaped transfer functions for ant lion optimization algorithm in feature selection problem. In: Proceedings of the International Conference on Future Networks and Distributed Systems, p. 14. ACM, July 2017
16. Anter, A., Ali, M.: Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems. Soft Comput. 1–20 (2019)
17. Anter, A.M., Azar, A.T., Fouad, K.M.: Intelligent hybrid approach for feature selection. In: Hassanien, A.E., Azar, A.T., Gaber, T., Bhatnagar, R., F. Tolba, M. (eds.) AMLTA 2019. AISC, vol. 921, pp. 71–79. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-14118-9_8
18. Rodrigues, D., et al.: A wrapper approach for feature selection based on bat algorithm and optimum-path forest. Expert Syst. Appl. **41**(5), 2250–2258 (2014)