JCSE Journal of Computing Science and Engineering

Regular Paper
Journal of Computing Science and Engineering,
Vol. 10, No. 4, December 2016, pp. 111-117

An ADHD Diagnostic Approach Based on Binary-Coded Genetic Algorithm and Extreme Learning Machine

Vasily Sachnev*
School of Information, Communication and Electronics Engineering, The Catholic University of Korea, Bucheon, Korea
bassvasys@hotmail.com

Sundaram Suresh
School of Computer Engineering, Nanyang Technological University, Singapore
ssundaram@ntu.edu.sg

Abstract
An accurate approach for diagnosis of attention deficit hyperactivity disorder (ADHD) is presented in this paper. The presented technique efficiently classifies three subtypes of ADHD (ADHD-C, ADHD-H, ADHD-I) and typically developing control (TDC) by using only structural magnetic resonance imaging (MRI). The research examines structural MRI of the hippocampus from the ADHD-200 database. Each available MRI has been processed by a region-of-interest (ROI) to build a set of features for further analysis. The presented ADHD diagnostic approach unifies feature selection and classification techniques. The feature selection technique based on the proposed binary-coded genetic algorithm searches for an optimal subset of features extracted from the hippocampus. The classification technique uses a chosen optimal subset of features for accurate classification of three subtypes of ADHD and TDC. In this study, the famous Extreme Learning Machine is used as a classification technique. Experimental results clearly indicate that the presented BCGA-ELM (binary-coded genetic algorithm coupled with Extreme Learning Machine) efficiently classifies TDC and three subtypes of ADHD and outperforms existing techniques.

Category: Smart and intelligent computing

Keywords: Attention deficit hyperactivity disorder; ADHD-200; Hippocampus; Binary-coded genetic algorithm; Extreme learning machine

I. INTRODUCTION

Attention deficit hyperactivity disorder (ADHD) is a neuropsychiatric disorder of children which affects around 5% of 7- to 21-year-old individuals. ADHD patients are usually hyperactive, abnormally impulsive, or both. According to recent studies, ADHD [1] has either a genetic or environmental nature, or both. However, the cause of ADHD is not fully understood [2].

ADHD research is mainly based on an analysis of magnetic resonance imaging (MRI). MRI makes clear images of the human brain and can be used to identify pathological parts. Recently researchers discovered new interesting effects which can lead us to a better understanding of the hidden mechanism of ADHD. For example, Ivanov et al. [3] found serious modifications in specific

Open Access http://dx.doi.org/10.5626/JCSE.2016.10.4.111 http://jcse.kiise.org

Received 01 November 2016; **Accepted** 05 December 2016
*Corresponding Author

 pISSN: 1976-4677 eISSN: 2093-8020

areas, i.e., the amygdala and the thalamus. Cherkasova and Hechtman [4] discovered a significant reduction in cerebral volume. Giedd and Rapoport [5] investigated different brain areas such as the frontostriatal areas, corpus callosum, basal ganglia, and temporoparietal lobes. Finally, it was proven that the amygdala, caudate, hippocampus, striatum, thalamus, basal ganglia, and some other brain areas are affected in diagnosed cases of ADHD.

A lack of emotion is a key symptomatic component of an ADHD diagnosis. The amygdala, caudate and hippocampus are brain regions that have been intensively studied in the medical literature in order to understand human emotions [6, 7]. The amygdala is responsible for emotional learning; the caudate region forms emotional memory, goal-orientation, and other forms of emotion; and the hippocampus is responsible for decision-making. Researchers have intensively studied the hippocampus for a better understanding of human emotions. Recent studies [8, 9] discovered significant modification in the hippocampus due to ADHD.

In this study, features extracted from the hippocampus were chosen to build an efficient ADHD diagnostic approach which classifies normal persons and three subtypes of ADHD.

MRI is widely used to build efficient diagnostic schemes. Badu et al. [10] designed the projection based learning algorithm of a meta-cognitive radial basis function network with recursive feature elimination (PBL-McRBFN-RFE) to build a classifier for Parkinson's disease (PD) and also identified brain areas probably responsible for PD. Mahanand et al. [11] adapted PBL-McRBFN for ADHD diagnosis. Later, Rangarajan et al. [12] improved the efficiency of PBL-McRBFN for ADHD diagnosis.

In this paper, we proposed an efficient ADHD diagnostic approach based on a binary-coded genetic algorithm coupled with Extreme Learning Machine (BCGA-ELM). The proposed method extracts a set of voxels from the MRI of the hippocampus area taken from ADHD-200 [13], using the region-of-interest (ROI) method. In the proposed method each voxel from the hippocampus area is coded using binary coefficients: "1" means that the voxel is chosen and "0" means the voxel is skipped. The BCGA searches for an optimal set of features (voxels), which is used to build an efficient ADHD classifier. An extensive search for a set of voxels with appropriate performances requires a machine-learning technique with high computational speed and acceptable generalization ability. Extreme Learning Machine [14] is the best candidate for the fitness function in the presented ADHD diagnostic framework. ELM is computationally less intensive, has good generalization ability, and is efficient for solving classification problems with 4 classes, as in this case, typically developing control (TDC) vs. ADHD-C vs. ADHD-H vs. ADHD-I.

The paper is organized as follows: in Section I, a data set of ADHD-200 is introduced. In Section II, the framework of the proposed BCGA-ELM is presented. Section III presents the experimental results. Finally, Section IV concludes the paper.

II. ADHD-200

ADHD-200 [13] is a set of MRI scans for ADHD research. ADHD-200 contains MRI scans from 941 persons (581 normal and 360 ADHD) taken from 7 brain areas (amygdala, caudate, cerebellar vermis, corpus callosum, hippocampus, striatum, and thalamus). A number of necessary features (voxels) have been extracted from each brain region using a ROI method: amygdala (1,050 voxels), caudate (3,904 voxels), cerebellar vermis (6,358 voxels), corpus callosum (8,536 voxels), hippocampus (6,076 voxels), striatum (9,359), and thalamus (6,438 voxels). Finally, the ADHD-200 database unifies a total of 41,721 voxels.

The 941 participants are divided into 4 groups:
1. 581 TDC or normal persons without ADHD.
2. 210 ADHD-C, ADHD combined subtype.
3. 137 ADHD-I, ADHD predominantly inattentive subtype.
4. 13 ADHD-H, patients with ADHD predominantly inattentive subtype.

The 941 subjects were further divided into training and testing subsets of 770 and 171 samples, respectively. The training subgroup contains 487 TDC, 161 ADHD-C, 11 ADHD-H, and 111 ADHD-I. The testing subgroup contains 94 TDC, 49 ADHD-C, and 2 ADHD-H.

In this paper, a set of 6,076 voxels extracted from the Hippocampus area is used to build an efficient ADHD classifier and to process the data.

III. PROPOSED BCGA-ELM APPROACH FOR ADHD DIAGNOSIS

The proposed ADHD diagnosis approach efficiently classifies TDC and three subtypes of ADHD (i.e., ADHD-C, ADHD-I, ADHD-H). The method contains 3 major steps:

Step 1. Feature extraction,
Step 2. Feature selection, and
Step 3. ADHD classification (see Fig. 1).

The ADHD diagnosis approach starts by accumulating necessary information for further analysis (see block "Feature extraction" in Fig. 1). In this paper, a set of MRI scans of the hippocampus taken from 941 patients available in the ADHD-200 database is processed by ROI to identify a set of 6,072 voxels (i.e., a 3D subspace in the human brain). The set of extracted voxels is then processed in the proposed "Feature selection" technique based on a BCGA. The "Feature selection" technique searches the

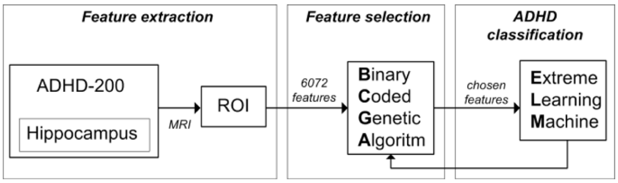

Fig. 1. The framework of the proposed attention deficit hyperactivity disorder (ADHD) diagnosis based on BCGA-ELM (binary-coded genetic algorithm coupled with Extreme Learning Machine).

set of best features with a promising performance for ADHD diagnosis. The chosen features are then utilized to build a 4-class ADHD classifier based on the ELM.

In the proposed ADHD diagnosis scheme the BCGA has been combined with the Extreme Learning Machine. The efficiency of the ADHD classification mostly depends on the chosen features. Sets of chosen features which produce better accuracy in ADHD diagnosis are used again to generate other, slightly different sets of features with even better accuracy in ADHD diagnosis. Sets of chosen features which produce low accuracy in ADHD diagnosis are discarded.

The detailed explanations of the proposed ADHD diagnosis approach are presented below.

A. Region-of-Interest

ROI creates a set of voxels from MRI scans available in the ADHD-200 database. Each MRI from the database is processed by the Burner pipeline obtained from the ADHD-200 consortium. The burner pipeline consists of three major steps. The first step utilizes statistical parametric mapping (SPM) [15] to segment all MRI into gray matter and white matter. In the second step, the tissue maps are normalized using DARTEL (diffeomorphic anatomical registration through exponentiated Lie algebra) [16]. In the third step, all images are iteratively assigned group templates and population averages. Finally, each image is transformed into a set of population averages.

In this research, ROI masks generated by using the pickatlas tool [17] were combined with data extracted from ADHD-200. ROI masks highlight the brain regions responsible for ADHD. In the final stage, a set of 6,072 features from the hippocampus area was extracted for further analysis.

B. Binary-Coded Genetic Algorithm

The BCGA is a variation of the famous genetic algorithm (GA) adapted for binary data. The GA utilizes self-adaptive mechanisms similar to those found in nature to search for optimal solutions for various scientific and engineering problems. The GA is a heuristic iterative optimization technique for optimization problems with an extremely large number of variables. GA consists of five major functional units: problem specification, genetic operators, fitness function, selection procedure, and termination criteria.

Problem specification: Any problem in GA should be presented as a set of relevant variables. The set of variables describes the solution for a given problem. Each solution is then evaluated by a fitness function. The number of possible solutions for problems commonly solvable by GA is extremely large, and so a search for an optimal solution becomes a serious challenge. A proper choice of the "problem specification" may simplify the search and speed up the necessary calculations.

In the proposed BCGA, each solution is a set of binary coefficients which represent the status of each feature/voxel in the hippocampus brain area (see Fig. 2). Each binary coefficient links to a corresponding feature from the hippocampus brain area. If the binary coefficient is "1" then the feature is selected; if "0" then the feature is skipped. Thus, the set of 6072 binary coefficients specifies a set of chosen features from the Hippocampus brain area (see Fig. 2).

Genetic operators: Crossovers and mutations are the two basic genetic operators in GA framework (see Fig. 3). Crossovers and mutations mimic natural genetic recombination. The crossover uses genetic material from two sources to create new combinations of genes, which unify genetic material from both sources. The new genetic combinations may or may not result in better properties. A mutation modifies genes randomly and may cause significant degradation of property. The mutation may also generate a new property which did not exist in an initial source. The combination of crossover and mutation is a key component of life in nature. Implementation of crossover and mutation in a single framework forms the basis for the GA.

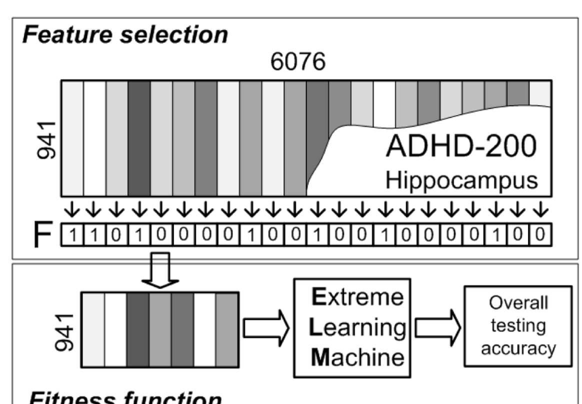

Fig. 2. The binary solution for attention deficit hyperactivity disorder (ADHD) diagnosis approaches based on BCGA-ELM (binary-coded genetic algorithm coupled with Extreme Learning Machine).

In the GA framework, the crossover operator picks two source solutions and creates a new solution using some predefined rules. Mutation in the GA framework modifies genes randomly.

The efficiency of the GA mostly depends on the design of the genetic operators (crossover and mutation). The design of the crossover and mutation is problem-specific and depends on the given data. Usually, researchers choose a crossover operator already present in the literature. As long as the performance of the GA highly depends on the optimization problem and given data, the choice of the proper crossover is a big challenge. One of the best strategies here can be the use of a hybrid crossover, where one crossover will be chosen randomly from a list of many given crossovers. Such strategy helps maintain the convergence of the GA for most of the solvable optimization problems.

The proposed BCGA uses crossover and mutation adapted for binary coefficients such as single point crossover, 2 point crossover, uniform crossover, arithmetic crossover and bit inversion mutation. The proposed BCGA uses the concept of hybrid crossover, which randomly picks one crossover among 4 given crossovers (the aforementioned single point crossover, 2 point crossover, uniform crossover and arithmetic crossover).

The fitness function is a special procedure to evaluate the solutions created by crossover and mutation. The fitness function computes a numerical value for further analysis. Solutions with the best value (minimum or maximum as depending on the problem) may be an optimal or suboptimal solution for a given problem. Such solutions are selected as the candidates to be processed by crossover and mutation again to get even better solutions. Solutions with a "bad" value are usually ignored.

In the proposed BCGA-ELM framework, the ELM evaluates each binary solution. A binary solution links to the set of features from the hippocampus area. Then the set of chosen features is used to train the ELM classifier. Finally, the overall testing accuracy of the ELM classifier evaluated on each binary solution is computed (see Fig. 2).

The selection procedure determines the chance for each new solution to be chosen during next GA generation. The procedure sorts solutions according to their corresponding fitness values (i.e., overall testing accuracies). Solutions with a better fitness value have a higher chance to be chosen for crossover and mutation compared to solutions with worse fitness values.

The proposed BCGA-ELM uses a geometric ranking method [18] as a selection procedure. In the geometric ranking method, all solutions are sorted in descending order of fitness value (overall testing accuracy of the ELM classifier). The probability of any solution j being selected is calculated as follows:

$$P_j = q'(1-q)^{r_j-1} \tag{1}$$

where

$$q' = \frac{q}{1-(1-q)^N}$$

q' is a selection score, r_j is the rank of the j-th solution in the partially ordered set, and N is the population size. The detailed explanation of the geometric ranking method is given in [18]. In this research, the parameter q is chosen as 10^{-3}.

Termination criteria: The GA stops when there is no improvement in terms of overall testing accuracy during the last 50 generations.

BCGA-ELM framework: The proposed BCGA-ELM starts with an initialization step (see Fig. 3). An initial population of 200 randomly generated binary solutions F^0_i (i = 1,2,3,…,200) with 6,072 binary coefficients is created. Then each binary solution F^0_i is used to build a set of corresponding features from the Hippocampus brain area (see "Feature selection" in Fig. 2). Chosen features are used to build the ELM classifier. The overall testing efficiency of ELM classifier gives the fitness value f^0_i. Thus, for each binary solution F^0_i the corresponding fitness value f^0_i is computed. The selection procedure processes all binary solutions and builds the initial population for BCGA (see "Initial population" in Fig. 3).

The GA iteratively updates populations by using crossover, mutation, ELM (fitness function) and selection procedures until the termination criteria is satisfied (see Fig. 3). In each population crossover creates 70% or 140 new solutions, mutation creates the remaining 30% or 60 new solutions. The Selection procedure then determines the chance of each solution from new population to be chosen during next generation.

C. Extreme Learning Machine

The proposed ADHD diagnosis approach based upon MRI from ADHD-200 data set efficiently solves the 4-class classification problem, i.e., ADHD-C vs. ADHD-H vs. ADHD-I vs. TDC. The problem has a high-dimensional feature space (6,072 features).

Any classifier is designed to approximate the functional

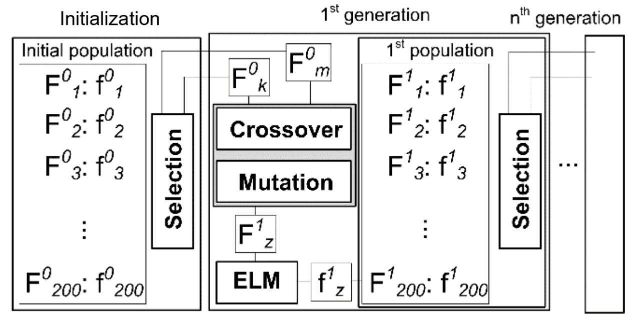

Fig. 3. The framework of the proposed BCGA-ELM (binary-coded genetic algorithm coupled with Extreme Learning Machine).

relationship between given features and class labels. In this paper, an ELM with hidden neurons controlled by a Gaussian activation function is used to approximate a decision boundary. ELM is a single hidden layer feed-forward neural network where input weights are randomly assigned, and output weights are estimated analytically [14]. The hidden neurons have randomly assigned bias.

The classification problem for ELM is defined as follows: the training data is a set of *N* samples, $\{(X^1, c^1), ..., (X^t, c^t), .., (X^N, c^N)\}$, where X^t is a vector of *m*-dimensional input features of *t*-th sample and $c^t \in \{1, 2, 3, ..., C\}$ is a class label. The coded class label y^t is calculated as follows:

$$y_k^t = \begin{cases} 1, & \text{if } c^t = k \\ -1, & \text{otherwise} \end{cases} \quad k = 1, 2, ..., C$$

where *C* is the number of classes. In our case, $C = 4$.

The classification problem approximates a decision function F_{ELM} and maps the input features to the coded class labels, i.e., $F_{ELM}: X \rightarrow y$. The input layer neurons in the ELM framework are linear, while the hidden layer neurons use a Gaussian activation function. ELM solves the problem with *m* input neurons and *C* output neurons. For further details about ELM, please refer to [14].

ELM has a significant drawback. The performance of the ELM mostly depends on the randomly chosen centers and hidden neuron biases, especially in applications like the ADHD-classification problem. Thus, in this method, a 10-fold validation approach is used to balance ELM performance. The ELM classifier is trained 10 times using the same features and different random-input neuron centers and hidden neuron biases.

IV. EXPERIMENTAL RESULTS

The proposed BCGA-ELM for ADHD diagnosis has been tested by diagnosing the 171 patients available in ADHD-200 data set. Classification performance of the best ELM classifier and set of voxels chosen by BCGA is presented in this section.

In this work, the concept of confusion matrices was utilized for a more accurate analysis. A confusion matrix (see Tables 1, 2) presents information about each examined sample according to the predicted class label (calculated by a machine-learning technique) in respect of the actual label. Table 1 is a "Training confusion matrix" and Table 2 is a "Testing confusion matrix". Both matrices are 4×4 in size. The vertical index represents the actual class label; the horizontal index represents the predicted class label. For example, in the "Training confusion matrix", the first column presents the distribution of all TDC samples according to the predicted class label. Two hundred seventy-four TDC samples among 487 were classified correctly (actual class label is the same with predicted class label), a total of 96 TDC samples were classified as ADHD-C (predicted class label links to ADHD-C), a total of 94 TDC samples were classified as ADHD-I (predicted class label links to ADHD-I), and 23 TDC were classified as ADHD-H. A similar analysis can be done for ADHD-C, ADHD-H, and ADHD-I samples. Thus, the confusion matrix shows the distribution of each examined sample according to its predicted class label and accurately portrays the performance of ADHD classification.

In this paper, overall training and testing efficiencies have been used to evaluate the performance of the proposed ADHD diagnosis approach. Overall testing η^{test} and training η^{train} accuracies are calculated as follows:

$$\eta^{test} = \frac{1}{N_{test}} \sum_{i=1}^{4} S_i^{test} \times 100\%$$

$$\eta^{train} = \frac{1}{N_{train}} \sum_{i=1}^{4} S_i^{train} \times 100\%$$

where S_i^{test} and S_i^{train} are numbers of correctly classified samples in class *i* for training and testing sets respectively; N_{test} and N_{train} are total numbers of samples available for training and testing.

Overall η^{test} and training η^{train} accuracies can be calculated using the confusion matrices. Assume $\boldsymbol{m}^{test}$ and $\boldsymbol{m}^{train}$ are confusion matrices for training and testing, then:

$$\eta^{test} = \frac{\sum_{i=1}^{4} \boldsymbol{m}_{ii}^{test}}{\sum_{i=1}^{4} \sum_{j=1}^{4} \boldsymbol{m}_{ij}^{test}}$$

$$\eta^{train} = \frac{\sum_{i=1}^{4} \boldsymbol{m}_{ii}^{train}}{\sum_{i=1}^{4} \sum_{j=1}^{4} \boldsymbol{m}_{ij}^{train}}$$

Table 1. Training confusion matrix

	TDC	ADHD-C	ADHD-H	ADHD-I
TDC	274	96	23	94
ADHD-C	28	109	5	19
ADHD-H	0	0	11	0
ADHD-I	20	9	1	81

ADHD: attention deficit hyperactivity disorder, TDC: typically developing control.

Table 2. Testing confusion matrix

	TDC	ADHD-C	ADHD-H	ADHD-I
TDC	63	15	1	15
ADHD-C	12	25	2	10
ADHD-H	1	0	0	1
ADHD-I	10	4	0	12

ADHD: attention deficit hyperactivity disorder, TDC: typically developing control.

Overall training accuracy is $\eta^{train} = 61.69\%$.

Overall testing accuracy is $\eta^{test} = 58.48\%$.

BCGA-ELM chooses 104 voxels from the hippocampus area to build the best ELM classifier.

A. Comparison with Existing Methods

Most researchers are trying to classify TDC vs. ADHD, which is a 2-class classification problem (or binary classification problem). The classification of the three ADHD subtypes (ADHD-I, ADHD-H, ADHD-C) and TDC, or the 4-class classification problem is a relatively new research area. Finding solutions for the TDC vs. ADHD-I vs. ADHD-H vs. ADHD-C classification problem is more challenging compared to TDC vs. ADHD and requires more sophisticated machine learning techniques to build an efficient classifier. The proposed BCGA-ELM is designed to solve the TDC vs. ADHD-I vs. ADHD-H vs. ADHD-C classification problem with acceptable accuracy.

A comparison with existing methods should cover ADHD diagnosis schemes focused on solving TDC vs. ADHD-I vs. ADHD-H vs. ADHD-C diagnosis using the complete set of 941 samples from the ADHD-200 data set. Sachnev [19] combined a similar BCGA with metacognitive neuro-fuzzy interface system (McFIS) to build an ADHD classifier in the hippocampus area. The authors used the suggested set of 770 samples for training and 171 samples for testing. Overall testing efficiency was 56%.

Kuang et al. [20] dealt with the prefrontal cortex, visual cortex and cingulate cortex brain areas and the 4-class classification problem for a reduced number of ADHD patients. The authors used a famous deep-learning approach to build a classifier. The authors reported an average testing accuracy of 35.19%.

Qureshi et al. [21] proposed an ELM-based classifier for 3 classes (TDC, ADHD-I, ADHD-C) and a reduced number of ADHD patients (159 patients). The authors reported a testing accuracy of 60.78%.

A direct comparison is valid only with the method presented by Sachnev [19]. The method proposed here shows a 2.48% improvement in terms of overall testing accuracy. A comparison with other techniques may not be valid. Both examined methods [20] and [21] use a reduced number of samples for training and testing. Experiments with the complete set of samples increase the complexity of the classifiers and reduce accuracy. The method presented in Qureshi et al. [21] solves the 3-class classification problem, which is less challenging, compared to the 4-class TDC vs. ADHD-I vs. ADHD-H vs. ADHD-C classification problem considered here.

V. CONCLUSION

An ADHD diagnosis approach for classifying TDC and three subtypes of ADHD is presented in this paper. The presented approach is based on the proposed BCGA-ELM algorithm, and efficiently solves the 4-class classification problem of TDC vs. ADHD-I vs. ADHD-H vs. ADHD-C. The BCGA searches an optimal set of voxels/features in the hippocampus brain area. The set of chosen feature/voxels is used to build an efficient ADHD classifier based on the ELM. Experimental results indicate clear performance advantage of the proposed method over the existing ADHD diagnosis approaches for which direct comparison is valid.

ACKNOWLEDGMENTS

This work was supported by The Catholic University of Korea research fund in 2016.

REFERENCES

1. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR)*, 4th ed., Washington, DC: American Psychiatric Association, 2000.
2. T. Banaschewski, K. Becker, S. Scherag, B. Franke, D. Coghill, "Molecular genetics of attention-deficit/hyperactivity disorder," *European Child and Adolescent Psychiatry*, vol. 19, no. 3, pp. 237-257, 2010.
3. I. Ivanov, R. Bansal, X. Hao, H. Zhu, C. Kellendonk, L. Miller, J. Sanchez-Pena, A. M. Miller, M. M. Chakravarty, K. Klahr, et al., "Morphological abnormalities of the thalamus in youths with attention deficit hyperactivity disorder," *American Journal of Psychiatry*, vol. 167, no. 4, pp. 397-408, 2010.
4. M. V. Cherkasova and L. Hechtman, "Neuroimaging in attention deficit hyperactivity disorder: beyond the frontostriatal circuitry," *Canadian Journal of Psychiatry*, vol. 54 no. 10, pp. 651-664, 2009.
5. J. N. Giedd and J. L. Rapoport, "Structural MRI of pediatric brain development: what have we learned and where are we going?," *Neuron*, vol. 67, no. 5, pp. 728-734, 2010.
6. H. Suzuki, K. Botteron, J. Luby, A. Belden, M. Gaffrey, C. Babb, T. Nishino, M. Miller, J. Ratnanather, and D. Barch, "Structural-functional correlations between hippocampal volume and cortico-limbic emotional responses in depressed children," *Cognitive, Affective, & Behavioral Neuroscience,* vol. 13, no. 1, pp. 135-151, 2013.
7. V. Vuontela, S. Carlson, A. M. Troberg, T. Fontell, P. Simola, S. Saarinen, and E. T. Aronen, "Working memory, attention, inhibition, and their relation to adaptive functioning and behavioral/emotional symptoms in school-aged children," *Child Psychiatry & Human Development*, vol. 44, no. 1, pp. 105-122, 2013.
8. M. H. Onnink, M. P. Zwiers, M. Hoogman, J. C. Mostert, C. C. Kan, J. Buitelaar, and B. Franke, "Brain alterations in adult ADHD: effects of gender, treatment and comorbid depression," *European Neuropsychopharmacology*, vol. 24, no. 3, pp. 397-409, 2014.

9. E. Perlov, A. Philipsen, L. T. van Elst, D. Ebert, J. Henning, S. Maier, E. Bubl, and B. Hesslinger, "Hippocampus and amygdala morphology in adults with 695 attention-deficit hyperactivity disorder," *Journal of Psychiatry and Neuroscience*, vol. 33, no. 6, pp. 509-515, 2008.
10. G. S. Babu, S. Suresh, and B. S. Mahanand, "A novel PBL-McRBFN-RFE approach for identification of critical brain regions responsible for Parkinson's disease," *Expert Systems with Applications*, vol. 41 no. 2, pp. 478-488, 2014.
11. B. S. Mahanand, R. Savitha, and S. Suresh "Computer aided diagnosis of ADHD using brain magnetic resonance images," in *Proceedings of 26th Australasian Joint Conference on Artificial Intelligence*, Dunedin, New Zealand, 2013, pp. 386-395.
12. B. Rangarajan, S. Suresh, and B. S. Mahanand, "Identification of potential biomarkers in the hippocampus region for the diagnosis of ADHD using PBL-McRBFN approach," in *Proceedings of 13th International Conference on Control, Automation, Robotics, and Vision (ICARCV)*, Singapore, 2014, pp. 17-22.
13. M. P. Milham, D. Fair, M. Mennes, and S. H. Mostofsky, "The ADHD-200 Consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience," *Frontiers in System Neuroscience*, vol. 6, article no. 62, 2012.
14. G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 985-990, 2006.
15. K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. D. Penny, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, Amsterdam: Elsevier/Academic Press, 2007.
16. J. Ashburner, "A fast diffeomorphic image registration algorithm," *Neuroimage*, vol. 38, no. 1, pp. 95-113, 2007.
17. J. A. Maldjian, P. J. Laurienti, R. A. Kraft, and J. H. Burdette, "An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets," *Neuroimage*, vol. 19, no 3, pp. 1233-1239, 2003
18. S. Suresh, S. N. Omkar, V. Mani, and T. G. Prakash, "Lift coefficient prediction at a high angle of attack using recurrent neural network," *Aerospace Science and Technology*, vol. 7, no. 8, pp. 595-602, 2003.
19. V. Sachnev, "An efficient classification scheme for ADHD problem based on binary-coded genetic algorithm and McFIS," in *Proceedings of 2015 International Conference on Cognitive Computing and Information Processing (CCIP)*, Noida, India, 2015, pp. 1-6.
20. D. Kuang, X. Guo, X. An, Y. Zhao, and L. He, "Discrimination of ADHD based on fMRI data with deep belief network," in *Proceedings of 10th International Conference on Intelligent Computing in Bioinformatics (ICIC)*, Taiyuan, China, 2014, pp. 225-232.
21. M. N. I. Qureshi, B. Min, H. J. Jo, and B. Lee, "Multiclass classification for the differential diagnosis on the ADHD subtypes using recursive feature elimination and hierarchical extreme learning machine: structural MRI study," *PLoS One*, vol. 11, no. 8, article no. e0160697, 2016.

Vasily Sachnev

Vasily Sachnev received his B.S. and M.S. degrees in Electrical Engineering from the Komsomolsk-na-Amure State Technical University, Russia, in 2002 and 2004, respectively. He received Ph.D. degree in Multimedia Security at the Center of Information Security and Technology (CIST), Graduate School of Information Management and Security, Korea University, Seoul, Korea in 2009. He joined the department of Information, Communication and Electronics Engineering of the Catholic University of Korea in 2010, where he is currently working as an assistant professor. His research interests include multimedia security, steganography, steganalysis, machine learning and bioinformatics.

Suresh Sundaram

Suresh Sundaram received his B.E. degree in electrical and electronics engineering from Bharathiyar University in 1999, and M.E. (2001) and Ph.D. (2005) degrees in Aerospace Engineering from Indian Institute of Science Bangalore, India. He was post-doctoral researcher in School of Electrical Engineering, Nanyang Technological University, Singapore, from 2005–2007. Subsequently, he was selected as an ERCIM research fellow for the period of 2007–2008 and he spent valuable time in the project team PULSAR at INRIA Sophia-Antipolis, France. For a short period, he was working as Faculty at Industrial Engineering, Korea University, Seoul. Later, he was with Indian Institute of Technology at Delhi, as an Assistant Professor in Electrical Engineering from 2008–2009. From 2010, he is working as Associate Professor in School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include computational cognitive system, neural networks, intelligent control, medical image processing, mathematical optimization, and game theory.