Identification of potential biomarkers in the hippocampus region for the diagnosis of ADHD using PBL-McRBFN approach

Rangarajan, B.; Suresh, S.; Mahanand, B. S.

2014

Rangarajan, B., Suresh, S., & Mahanand, B. S. (2014). Identification of potential biomarkers in the hippocampus region for the diagnosis of ADHD using PBL-McRBFN approach. 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), 17-22.

https://hdl.handle.net/10356/107096

https://doi.org/10.1109/ICARCV.2014.7064272

Identification of Potential Biomarkers in the Hippocampus Region for the Diagnosis of ADHD using PBL-McRBFN Approach

Rangarajan B.
School of Computer Engineering
Nanyang Technological University
Singapore
Email: badrinarayanan@ntu.edu.sg

Suresh S.
School of Computer Engineering
Nanyang Technological University
Singapore
Email: ssundaram@ntu.edu.sg

Mahanand B. S.
Department of Information Science
and Engineering, Sri Jayachamarajendra
College of Engineering, India
Email: bsmahanand@sjce.ac.in

***Abstract*—Attention Deficiency Hyperactivity Disorder (ADHD) as a disruptive behavior disorder is receiving lots of attention because of its complexity and need for early detection. This paper presents a study on identification of potential biomarkers in the diagnosis of ADHD based on the structural-MRI of the brain obtained through ADHD-200 competition data set. The region of the brain considered here is "hippocampus". The grey matter probability of the T1 images is segmented followed by tissue alignment and inter subject normalization. Then, the voxels of the hippocampus are segregated using a region-of-interest mask, and the grey matter tissue probability values are obtained. These values are then used as features to classify ADHD patients against typically developing controls using a projection based learning algorithm for a meta-cognitive radial basis function network (PBL-McRBFN) and compared the results with that of support vector machines. Initially we take all the voxels of hippocampus for our study and then we have selected the most relevant voxels as a biomarker using Chi-square approach and developed a classifier to diagnosis ADHD. The results clearly highlight that use of hippocampus from the structural-MRI is sufficient to diagnosis ADHD to certain degree of confidence.**

***Index Terms*—Attention Deficient Hyperactivity Disorder, Hippocampus, Meta-cognitive Radial Basis Function Network, Projection Based Learning, Region of Interest, classification.**

I. Introduction

Attention Deficiency Hyperactivity Disorder (ADHD) [15] is a disruptive behavior disorder [7] that is comorbid with other neurological disorders [22], [30]. Nearly 5% of school-going children worldwide are affected by ADHD [26]. The recent criteria laid by the Diagnostic and Statistical Manual of Mental disorders, V edition (DSM-V) [31] defines ADHD as a pattern of hyperactivity and/or inattention which runs irrespective of age and gender. The complete cause of ADHD is still unknown due to various reasons like comorbidity of other neurological disorders, age-wise in-congeniality, genetics, environment, and gender [6].

There are several regions of the brain that get affected when a person is diagnosed with ADHD. Studies have been conducted to understand the neurobiology of ADHD in children [9]. One of the key regions is "hippocampus". During rational decision-making, studies show that people who have ADHD, exhibit less activation in the hippocampus [12]. This results in the abnormality of its volume [22]. ADHD patients possess larger hippocampus when compared to that of typically developing controls (TDC) [25]. Because of the importance of hippocampus in decision-making and its relation with ADHD (as explained above), we have considered "hippocampus" as a region-of-interest in this study.

Researchers are motivated to understand the pathophysiology of ADHD using structural Magnetic Resonance Imaging (MRI) for many years. Recent works in [11] show that the abnormalities in brain volumes can be a measure to the diagnosis of ADHD. Among various regions, the hippocampus is shown to have a larger role in the diagnosis of ADHD. Children with ADHD have shown to have a large hippocampus when compared to that of adults [32]. However, these findings are based on different data sets, using different tools for analysis. Although different regions are being studied, understanding the onset and development of ADHD is limited by the lack of studies that analyze the contribution of individual regions using a single data set.

We present a region-of-interest (ROI) based feature extraction technique for structural-MRI to identify the potential biomarkers in the diagnosis of ADHD. In this study, we tried to understand the variations in the hippocampus region of ADHD-200 competition dataset [20]. Researchers have analyzed the ADHD using functional-MRI due to its complexity [23]. Only a few of them have conducted their studies using structural-MRI [24], [18], [33].

In this study, we examine the voxels of the hippocampus that are extracted from only the structural-MRI of the ADHD-200 data set and classify them as either TDC or ADHD. The voxels are obtained by extracting the ROI that is defined using the Wake Forest University Pick-atlas [16], [19]. Then the voxels are classified using the Projection Based Learning algorithm of a Metacognitive Radial Basis Function Network Classifier (PBL-McRBFN) [5], [2] as either TDC or ADHD. We further employ the Chi2 algorithm based feature selection [17] to identify bio-markers within hippocampus. Chi2 is based on χ^2 (Chi-Square) statistic which measures the dependence

between stochastic variables by calculating the highest significance level for all the attributes. Then for every pair of attributes, ChiMerge (a part of Chi2) automatically increments the χ^2 threshold which results in ranking the most relevant features. Once we have the reduced feature set, we study the behavior of the features using PBL-McRBFN classifier and compare the results with that of support vector machines (SVM) [8].

In the literature of human psychology, it has been suggested that metacognition empowers the learner with a self-regulated learning mechanism. Metacognition provides a means to accurately assess ones current knowledge, identify when a new knowledge is needed, as well as provide strategies to acquire that new knowledge [39], [14]. The meta-cognitive learning algorithm has been implemented in neuro-fuzzy systems [36], [34], [35], in complex-valued networks [28], [27], in neural networks [2] and in optimization [37]. The McRBFN classifier reproduces the Nelson and Narens model (Refer figure 1) of human meta-cognition [21]. The self regulatory learning

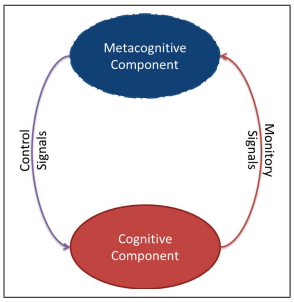

Fig. 1. Nelson and Narens Model Of Metacognition

mechanism of McRBFN uses the best human learning strategy, namely, what-to-learn, when-to-learn and how-to-learn in a meta-cognitive framework. Among various implementations, PBL-McRBFN [3], [4] is computationally efficient and provides better generalization performance.

Projection based learning algorithm of the PBL-McRBFN consists of a radial basis function network. The hidden layer has a Gaussian activation function which is the cognitive component of McRBFN. PBL-McRBFN also consists of a self-regulatory learning mechanism which is its meta-cognitive component. Comparing the performance of PBL-McRBFN and SVM proves that PBL-McRBFN has a better generalization in diagnosing ADHD by using the voxels of hippocampus and shows improved classification performance in higher dimensional data.

The paper is organized as follows: In Section II, the ADHD-200 data set is described, followed by the feature extraction technique and the description of the PBL-McRBFN classifier. Sections III-A and III-B present the results of study of TDC Vs ADHD using PBL-McRBFN and SVM, initially with all the voxels and then with the reduced set of voxels (Chi2 based). Finally, Section V presents the conclusions from the study.

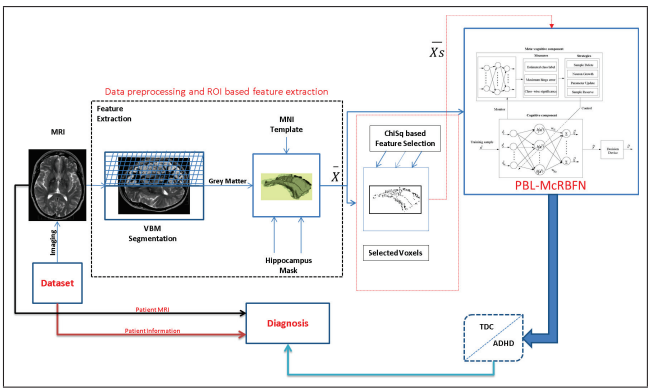

Fig. 2. Proposed automatic ADHD diagnostic mechanism using structural-MRI.

II. Methods and Materials

In this section, we present the complete details of the ADHD-200 data set, the ROI based feature extraction, the PBL-McRBFN classifier and proposed ADHD diagnosis using PBL-McRBFN. Fig. 2 presents the pictorial representation of the proposed automatic ADHD diagnostic mechanism using structural-MRI. The proposed approach takes the images (from MRI) of the subjects as input. The images are then processed for feature extraction. Here, the MNI template of hippocampus is used to extract its voxels, which are then sent to the classifier. Here we study the voxels using PBL-McRBFN and SVM as explained below.

- Initially, all the voxels ($\bar{X}$) of the hippocampus are classified as either ADHD or TDC.
- Then, few important voxels ($\bar{X}$s) is selected (using Chi2 based feature selection) and classified as either ADHD or TDC.

In the following subsections, the subcomponents of proposed automatic ADHD diagnostic mechanism is described in detail.

A. Data set

We use the ADHD-200 competition data set [20] in our study. This data set is a collection of brain MR images of 941 subjects from 8 participating members of the consortium.The subjects include 581 TDC and 360 ADHD as classes. Out of the 941 subjects, 770 subjects are provided as training data and 171 subjects are provided as testing data.

B. Data Preprocessing and ROI based feature extraction

All MRI data were processed with the Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) [1] based on the Statistical Parametric Mapping (SPM) [13] software package in the Burner pipeline from the ADHD-200 consortium. The Burner pipeline includes normalized grey matter maps generated using SPM. There are three steps in Burner pipeline. Firstly, T1 images are segmented into grey matter and white matter probability maps using SPM. Images are then iteratively registered to the group average which is updated iteratively. Lastly, the registration parameters are applied to each image to transform each image into the

space of population average. Modulation is applied to conserve the global tissue volumes after normalization.

In this paper, region-of-interest (ROI) based feature extraction mechanism have been employed. Traditionally, the ROI based feature extraction is used to calculate the anatomical volumes and also to investigate abnormal tissue structures. The functional significance of hippocampus includes formation of new memories, detection of novel events, factual memory and navigational presence. The ROI based feature extraction process is shown in Fig. 2. From the figure, it can be seen that the probability of grey matter tissue is extracted as features from all the modulated and normalized grey matter images, using the generated ROI masks. These ROIs are defined using the Wake Forest University Pickatlas [16], [19].

C. Projection Based Learning Algorithm for a Meta-cognitive Radial Basis Function Network

In this paper, we use the projection based learning algorithm for a meta-cognitive radial basis function network [2] classifier to distinguish the ADHD patients from TDC. In this section, we explain the learning algorithm of the classifier in detail.

Given the training data set with N samples, $(\mathbf{x}^1, c^1), \cdots, (\mathbf{x}^t, c^t), \cdots, (\mathbf{x}^N, c^N)$, where $\mathbf{x}^t = [x_1^t, \cdots, x_m^t]^T \in \Re^m$ is the m-dimensional input of the t^{th} sample, and c^t is its class label, the coded class labels $\left(\mathbf{y}^t = \left[y_1^t, \cdots, y_j^t, \cdots, y_n^t\right]^T\right) \in \Re^n$ are obtained as:

$$y_j^t = \begin{cases} 1 & \text{if} \quad c^t = j \\ -1 & \text{otherwise} \end{cases} \quad j = 1, \cdots n; \; t = 1, \cdots N \tag{1}$$

where n is the total number of classes. Approximating the decision function that maps $\mathbf{x}^t \in \Re^m \rightarrow \mathbf{y}^t \in \Re^n$ is the main objective of McRBFN. The classifier iterates the samples starting with zero hidden neurons and applies suitable strategy for the chosen sample. The learning algorithm and architecture of PBL-McRBFN are as follows:

1) McRBFN Architecture: There are two components of McRBFN; cognitive and meta-cognitive component. Fig. 3 shows the schematic diagram of the McRBFN classifier. The cognitive component possesses a feed-forward three-layered RBF network while the meta-cognitive component possesses a self-regulating model of the cognitive component. The relative knowledge is calculated by the meta-cognitive component whenever a new training sample is presented. Learning is controlled by choosing suitable strategies that addresses *what-to-learn*, *when-to-learn* and *how-to-learn* for the current sample based on this judgement of knowledge. We further explain two components of McRBFN in the following sections:

Cognitive component of McRBFN: The cognitive component of McRBFN is a feed-forward three-layered radial basis function network with linear input and output layers. It employs the gaussian activation function in its hidden layers.

The McRBFN employs K gaussian neurons from $t-1$ training samples. The input sample is given by $\mathbf{x}^t$ and predicted output of the j^{th} output neuron ($\widehat{y}_j^t$) of McRBFN is

$$\widehat{y}_j^t \quad = \quad \sum_{k=1}^{K} w_{kj} h_k^t, \qquad j = 1, 2, \cdots, n \tag{2}$$

where w_{kj} is the weight connecting the k^{th} hidden neuron to the j^{th} output neuron and h_k^t is the response of the k^{th} hidden neuron to the input $\mathbf{x}^t$ is given by

$$h_k^t \quad = \quad \exp\left(-\frac{\|\mathbf{x}^t - \boldsymbol{\mu}_k^l\|^2}{(\sigma_k^l)^2}\right) \tag{3}$$

where $\boldsymbol{\mu}_k^l \in \Re^m$ is the center and $\sigma_k^l \in \Re^+$ is the width of the k^{th} hidden neuron. The class to which the neuron k is represented by the superscript l.

During learning, a projection based learning (PBL) algorithm is used by the cognitive component. The following section describes the PBL algorithm in detail:

Projection based learning algorithm: The projection based learning algorithm calculates the network output parameters by minimizing the energy function.

For t training samples, the overall energy function is given by

$$J(\mathbf{W}) = \frac{1}{2} \sum_{i=1}^{t} \sum_{j=1}^{n} \left(y_j^i - \sum_{k=1}^{K} w_{kj} h_k^i\right)^2 \tag{4}$$

where h_k^i is the response of the k^{th} hidden neuron for i^{th} training sample.

The optimal output weights ($\mathbf{W}^* \in \Re^{K \times n}$) are estimated such that the total energy reaches its minimum.

$$\mathbf{W}^* \quad := \quad arg \min_{\mathbf{W} \in \Re^{K \times n}} \quad J(\mathbf{W}) \tag{5}$$

Let $\mathbf{W}^* \in \Re^{K \times n}$, then $\mathbf{W}^*$ is the optimal output weight corresponding to the minimum of the energy function if $J(\mathbf{W}^*) \leq J(\mathbf{W}) \; \forall \; \mathbf{W} \in \Re^{K \times n}$. The solution for the system of equations in Eq. (4) and (5) can be reduced in the form as follows:

$$\mathbf{W}^* \quad = \quad \mathbf{A}^{-1}\mathbf{B} \tag{6}$$

where A is the projection matrix and B is the output matrix. Please refer to [2] for complete details of the initialization and thresholds of meta-cognition.

Meta-cognitive component of McRBFN: The meta-cognitive component uses several measures of knowledge in the new training sample and uses the knowledge to control the cognitive component. It includes estimated class label ($\widehat{c}^t$), maximum hinge error (E^t) and spherical potential. These measures are further explained below: *Estimated Class label ($\widehat{c}^t$)*: Using the predicted output ($\widehat{\mathbf{y}}^t$), the estimated class label ($\widehat{c}^t$) can be obtained as

$$\widehat{c}^t = arg \max_{j \in 1, 2, \cdots, n} \widehat{y}_j^t \tag{7}$$

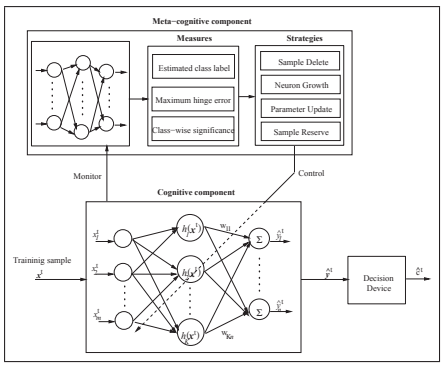

Fig. 3. Schematic diagram of McRBFN Classifier.

A class specific spherical potential is used to measure significance of sample, which is given in Eqn. (8),

$$\psi \approx -\frac{2}{K}\sum_{k=1}^{K} h\left(\mathbf{x}^t, \boldsymbol{\mu}_k^l\right) \tag{8}$$

2) Learning Strategies: The meta-cognitive component creates various strategies for learning based on the measures explained above. It supports the fundamental nature of self-regulated human-like learning (i.e., *what-to-learn*, *when-to-learn* and *how-to-learn*). The learning process in the cognitive component is controlled by the meta-cognitive component. The meta-cognitive component selects one of the following four learning strategies for every new training sample.

- *Sample Delete Strategy*: Delete strategy is used to check if the knowledge of the new sample is similar to the one that is present in the cognitive component. This strategy deletes the redundant sample.
- *Neuron Growth Strategy*: Neuron growth strategy is used to add a hidden neuron in the cognitive component.
- *Parameter Update Strategy*: Parameter update strategy updates the parameters whenever a new training sample is chosen.
- *Sample Reserve Strategy*: Finally, the sample reserve strategy is used to save some samples depending on the credibility of the information. If it contains some insignificant information, it can better be used at a later stage for fine tuning the parameters.

It must be noted that while the sample delete strategy addresses *what-to-learn*, the neuron addition and the parameter update strategies address *how-to-learn* components of self-regulation. For complete details of the initialization of the various thresholds of meta-cognition, one must refer to [2].

In the next section, we present the performance results of the automatic diagnosis of ADHD Vs TDC,

III. Experimental Results

In this section, we study the performance of the PBL-McRBFN classifier in distinguishing ADHD patients from typically developing controls. According to the ADHD-200 competition data set, 770 samples are chosen to train the classifier, and 171 samples are used in testing. Now, we evaluate the performance of PBL-McRBFN classifier using the overall efficiency and *F1* score:

Overall classification efficiency: The overall classification efficiency (η_o) is given as

$$\eta_o \quad = \quad \frac{\sum_{l=1}^{n} q_{ll}}{N} \times 100\% \tag{9}$$

where q_{ll} is the total number of correctly classified samples in the training/testing data set.

***F1* Score:** The *F1* score [10] is given as

$$F1 = 2 * (precision * recall)/(precision + recall). \tag{10}$$

A. Results using complete features

In this section, we present the diagnostic results of PBL-McRBFN classifier compared with SVM for all the voxels in classifying the subjects as TDC or ADHD patients. The number of voxels in each of the selected regions obtained by ROI selection, the number of neurons, the training overall efficiency and *F1* score, are tabulated in Table I. From the table, it can be observed that the overall efficiency of PBL-McRBFN is 67.84% in the testing set which is around 13% larger than that of SVM. Also, it can be observed that PBL-McRBFN uses smaller number of neurons while training.

TABLE I
Performance of PBL-McRBFN Vs SVM considering all the voxels of hippocampus.

Classifier	Voxels	K	Train		Test	
			η_o %	*F1*	η_o %	*F1*
McRBFN	6072	351	91.56	0.93	67.84	0.72
SVM	6072	632	83.09	0.82	54.38	0.53

The number of voxels in hippocampus region is 6072. Not all of them would contribute towards the diagnosis of ADHD. There is a need to select a subset of voxels as biomarkers. To achieve this, we have conducted a similar experiment with reduced set of features in the following section.

B. Identifying biomarkers in hippocampus

In this section, we identify the image biomarkers within hippocampus for classification. All the voxels within the selected regions of the brain may not contribute to the onset and development of ADHD. Here, we employ Chi2 based feature selection method to select few relevant voxels (as biomarkers). Then we use PBL-McRBFN and SVM to classify the selected voxels. Table II presents the results from this study. It can be observed from the table that the Chi2 method identified only 648 among 6042 voxels in the hippocampus region. The overall efficiency of PBL-McRBFN is 70.18% and is around 18% more than that of SVM. Also, it can be observed that PBL-McRBFN uses smaller number of neurons while training.

It can be observed from the tables (I and II) that the *F1* score of PBL-McRFBN classifier is 19% more using all the voxels (6072) and 26% more using the reduced set of voxels (648) which explains that PBL-McRBFN has better retrieval when compared to SVM. This proves the better

TABLE II
PERFORMANCE OF PBL-MCRBFN VS SVM CONSIDERING THE VOXELS CHOSEN USING CHI2.

Classifier	Voxels	K	Train		Test	
			η_o %	*F1*	η_o %	*F1*
McRBFN	648	347	88.70	0.90	70.18	0.72
SVM	648	649	80.23	0.78	52.04	0.46

generalization capability of PBL-McRBFN in classifying the voxels of hippocampus. The testing efficiency of 70.18% with an F1 score of 0.72 (Refer table II) shows a way that ADHD can be diagnosed using a subset of voxels of the hippocampus region.

We extend our analysis by studying a subset of the best ranked voxels using Chi2 based feature selection. Chi2 algorithm ranks the features based on the stochastic relevance of the pairs of attributes. We ranked the voxels of hippocampus and classified them further using PBL-McRBFN. Figure 4 shows the performance study of PBL-McRBFN along with the number of voxels ranging from 10 to 80 chosen based on the Chi2 algorithm. The x-axis depicts the number of features chosen, while the left y-axis shows the percentage of testing average and the right y-axis shows the F1 score. It can be observed that choosing the best 60 voxels of hippocampus gives an efficiency of around 72.5%. These 60 voxels of hippocampus could be considered as potential biomarkers in classifying ADHD. It can also be observed that the efficiency increases as the number of features increases from 10 to 60 and then it saturates. But, there is a steady increase in the F1 score as the number of features increases. This again proves a better generalization and retrieval of PBL-McRBFN.

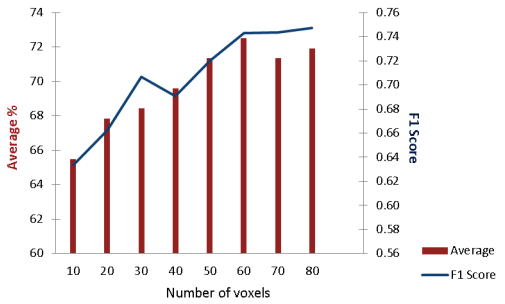

Fig. 4. Performance of PBL-McRBFN considering a subset of the best ranked voxels using Chi2.

IV. DISCUSSION

Efficient analysis of ADHD requires representative biomarkers (features) as well as a classifier which can learn to approximate the decision surface from the given features. Identification of biomarkers has been studied for many years in the medical literature with the focus on neurological and psychological disorders including ADHD [38], [29]. Recently, the radiological society of North America has published an article on "MRI Technique May Help Prevent ADHD Misdiagnosis" which stresses the fact the diagnosis of ADHD requires the identification of potential biomarkers in MRI. In this study, we focus on the identification of potential biomarkers for the diagnosis of ADHD using hippocampus as the region of interest. The identification of biomarkers is carried out using Chi2 based feature selection technique. We have obtained an overall testing accuracy of 70.18% choosing 648 voxels (features) from hippocampus. In addition to an efficient feature representation technique, a good learning mechanism to learn the decision surface separating the ADHD from TDC is mandatory. In literature, support vector machine (SVM) based approaches have been widely used for solving problems in the domain of medical analytics. However, the recent literature on soft computing approaches has shown that PBL-McRBFN based classification technique can achieve the better generalization ability in comparison to SVM. Further, PBL algorithm solves a system of linear equations where as SVM solves a quadratic optimization. This results in the reduction of computational cost when compared to that of SVM. PBL-McRBFN can estimate posterior probabilities more accurately than hinge-loss error or mean-squared error based loss functions. SVM uses all the samples for learning which results in over-fitting. PBL-McRBFN with its ability to self-regulate helps the networks build an efficient classifier. The above facts have motivated us to employ such a learning mechanism to efficiently diagnose ADHD based on Chi2 based features. In this study, we have used the state-of-the-art machine learning technique (PBL-McRBFN) to classify the test samples (from ADHD-200 competition data set) and compare the result with the a well-known classifier in literature (SVM in this case). The extended analysis is made to show that choosing a subset of voxels as biomarkers (from 10 to 80 in this case) help to better understand the classification accuracy of PBL-McRBFN. We would like to take this study forward by taking other regions individually and combined together to understand the diagnosis of ADHD as a whole.

V. CONCLUSION

This study proposes an identification of potential biomarkers in the diagnosis of Attention Deficiency Hyperactivity Disorder based on the structural-MRI of the brain obtained through the ADHD-200 competition data set. The hippocampus, which plays a key role in the formation of new neurons in the brain and due to its complexity in its volume is taken as the region-of-interest. The probability values of the grey matter, tissues of the hippocampus are segmented from the MR images. These values are then used as the features to classify ADHD patients against typically developing controls using PBL-McRBFN and SVM. Initially, all the 6072 voxels from hippocampus are considered. Then, using Chi2 based feature selection method, we studied the 648 voxels. We further extended the analysis by creating a subset of the best ranked voxels ranging from 10 to 80 and compared its performance using PBL-McRBFN. In the analysis, we found that ADHD can be diagnosed using a subset of voxels selected from the hippocampus region taken only from the structural-MRI. The performance of PBL-McRBFN is better in generalization and

in classification efficiency for reduced number of hippocampus voxels than that of SVM. It is observed in our study that considering only 60 voxels of hippocampus as potential biomarkers, we could able to classify ADHD against TDC with an F1 score of 0.74 and an efficiency close to 72.5%. We would like to extend this study further to classify the ADHD-200 data using structural-MRI for all the other regions.

VI. Acknowledgement

We would like to thank the ADHD-200 consortium for making the MRI data publicly available. The first and second authors would like to gratefully extend their thanks to the funding agency 'Nanyang Technological University - Ministry of Defence, Singapore (Grant No: MINDEF-NTU-JPP/13/02/01)' for their financial support in making this study possible.

References

[1] J. Ashburner, "A fast diffeomorphic image registration algorithm," *NeuroImage*, vol. 38, no. 1, pp. 95–113, 2007.

[2] G. S. Babu and S. Suresh, "Meta-cognitive neural network for classification problems in a sequential learning framework," *Neurocomputing*, vol. 81, no. 0, pp. 86–96, 2012.

[3] ——, "Meta-cognitive RBF network and its based learning algorithm for classification problems," *Applied Soft Computing Journal*, vol. 13, no. 1, pp. 654–666, 2013.

[4] ——, "Sequential projection-based metacognitive learning in a radial basis function network for classification problems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 2, pp. 194–206, Feb 2013.

[5] 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.

[6] T. Banaschewski, K. Becker, S. Scherag, B. Franke, and D. Coghill, "Molecular genetics of attention-deficit hyperactivity disorder: An overview," *European Child and Adolescent Psychiatry*, vol. 19, no. 3, pp. 237–257, 2010.

[7] O. G. Bukstein, "Disruptive behavior disorders and substance use disorders in adolescents," *Journal of Psychoactive Drugs*, vol. 32, no. 1, pp. 67–79, 2000.

[8] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.

[9] S. Cortese, "The neurobiology and genetics of attention-deficit hyperactivity disorder: What every clinician should know," *European Journal of Paediatric Neurology*, 2012.

[10] M. W. David Powers, "Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation," School of Informatics and Engineering, Flinders University, Adelaide, Australia, Tech. Rep. SIE-07-001, 2007.

[11] A. De La Fuente, S. Xia, C. Branch, and X. Li, "A review of attention-deficit/hyperactivity disorder from the perspective of brain networks," *Frontiers in Human Neuroscience*, vol. 7, no. 192, 2013.

[12] M. Ernst, A. S. Kimes, E. D. London, J. A. Matochik, D. Eldreth, S. Tata, C. Contoreggi, M. Leff, and K. Bolla, "Neural substrates of decision making in adults with attention deficit hyperactivity disorder," *American Journal of Psychiatry*, vol. 160, no. 6, pp. 1061–1070, 2003.

[13] K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. Penny, Eds., *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, 2007.

[14] R. Isaacson and F. Fujita, "Metacognitive knowledge monitoring and self-regulated learning: Academic success and reflections on learning," *Journal of the Scholarship of Teaching and Learning*, vol. 6, no. 1, pp. 39–55, 2006.

[15] A. Krain and F. Castellanos, "Brain development and ADHD," *Clinical Psychology Review*, vol. 26, no. 4, pp. 433–444, 2006.

[16] J. L. Lancaster, M. G. Woldorff, L. M. Parsons, M. Liotti, C. S. Freitas, L. Rainey, P. V. Kochunov, D. Nickerson, S. A. Mikiten, and P. T. Fox, "Automated talairach atlas labels for functional brain mapping," *Human Brain Mapping*, 2000.

[17] H. Liu and R. Setiono, "Chi2: feature selection and discretization of numeric attributes," in *Tools with Artificial Intelligence, 1995. Proceedings., Seventh International Conference on*, Nov 1995, pp. 388–391.

[18] B. S. Mahanand, R. Savitha, and S. Suresh, *Computer Aided Diagnosis of ADHD Using Brain Magnetic Resonance Images*, ser. Lecture Notes in Computer Science. Springer International Publishing, 2013, vol. 8272, pp. 386–395.

[19] J. Maldjian, P. Laurienti, R. Kraft, and J. Burdette, "An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets," *NeuroImage*, 2003.

[20] 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 Systems Neuroscience*, vol. 6, no. 62, 2012.

[21] T. O. Nelson and L. Narens, "Metamemory: A theoretical framework and new findings," *Psychology of Learning and Motication*, vol. 26, no. C, pp. 125–173, 1990.

[22] A. 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, MAR 2014.

[23] Y. Paloyelis, M. Mehta, J. Kuntsi, and P. Asherson, "Functional magnetic resonance imaging in attention deficit hyperactivity disorder (ADHD): a systematic literature review," *Expert Review of Neurotherapeutics*, 2007.

[24] X. Peng, P. Lin, T. Zhang, and J. Wang, "Extreme learning machine-based classification of ADHD using brain structural MRI data," *PLoS ONE*, vol. 8, no. 11, 11 2013.

[25] K. Plessen, R. Bansal, H. Zhu, and et al, "Hippocampus and amygdala morphology in attention-deficit/hyperactivity disorder," *Archives of General Psychiatry*, vol. 63, no. 7, pp. 795–807, 2006.

[26] G. Polanczyk, M. S. de Lima, B. L. Horta, J. Biederman, and L. A. Rohde, "The worldwide prevalence of ADHD: a systematic review and metaregression analysis," *American Journal of Psychiatry*, vol. 164, no. 6, pp. 942–948, 2007.

[27] R. Savitha, S. Suresh, and N. Sundararajan, "A meta-cognitive learning algorithm for a fully complex-valued relaxation network," *Neural Networks*, vol. 32, pp. 209–218, 2012.

[28] ——, "Metacognitive learning in a fully complex-valued radial basis function neural network," *Neural Computation*, vol. 24, no. 5, pp. 1297–1328, 2012.

[29] C. Scassellati, C. Bonvicini, S. V. Faraone, and M. Gennarelli, "Biomarkers and attention-deficit hyperactivity disorder: A systematic review and meta-analyses," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 51, no. 10, pp. 1003–1019, 2012.

[30] W. T. Spencer T, Biederman J, "Attention-deficit/hyperactivity disorder and comorbidity," *Pediatric Clinics North America*, vol. 46, no. 5, pp. 915–27, 1995.

[31] S. Steinau, "Diagnostic criteria in attention deficit hyperactivity disorder - changes in DSM-5," *Frontiers in Psychiatry*, vol. 4, no. 49, 2013.

[32] M. C. Stevens and E. Haney-Caron, "Comparison of brain volume abnormalities between ADHD and conduct disorder in adolescence." *J. Psychiatry Neurosci.*, vol. 37, no. 6, pp. 389–98, Nov. 2012.

[33] K. Subramanian, R. Savitha, and S. Suresh, "A meta-cognitive interval type-2 fuzzy inference system classifier and its projection based learning algorithm," in *IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, April 2013, pp. 48–55.

[34] K. Subramanian and S. Suresh, "Human action recognition using meta-cognitive neuro-fuzzy inference system," *International Journal of Neural Systems*, vol. 22, no. 06, p. 1250028, 2012.

[35] ——, "A meta-cognitive sequential learning algorithm for neuro-fuzzy inference system," *Applied Soft Computing*, vol. 12, no. 11, pp. 3603–3614, 2012.

[36] K. Subramanian, S. Suresh, and N. Sundararajan, "A metacognitive neuro-fuzzy inference system (McFIS) for sequential classification problems," *IEEE Transactions on Fuzzy Systems*, vol. 21, no. 6, pp. 1080–1095, Dec 2013.

[37] M. R. Tanweer and S. Suresh, "Human cognition inspired particle swarm optimization algorithm," in *IEEE 9th Int. Conf. on Intelligent Sensors, Sensor Networks and Information Processing*, 2014, pp. 1–6.

[38] D. Wallis, "The search for biomarkers for attention deficit hyperactivity disorder," *Drug news and perspectives*, vol. 23, no. 7, September 2010.

[39] A. L. Wenden, "Metacognitive knowledge and language learning," *Applied Linguistics*, vol. 19, no. 4, pp. 515–537, 1998.