



Accurate detection of autism spectrum disorder from structural MRI using extended metacognitive radial basis function network



Vigneshwaran Subbaraju^a, Suresh Sundaram^{a,*}, Sundararajan Narasimhan^b, Mahanand Belathur Suresh^c

^a School of Computer Engineering, Nanyang Technological University, Singapore

^b School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

^c Department of Information Science and Engineering, Sri Jayachamarajendra College of Engineering, Mysore, India

ARTICLE INFO

Keywords:

Autism spectrum disorder
Magnetic resonance imaging
Voxel-based morphometry feature extraction
Metacognitive radial basis function network classifier
Projection based learning algorithm
q-Gaussian activation function

ABSTRACT

In this paper, we present an accurate detection of Autism Spectrum Disorder (ASD) from structural MRI using an Extended Metacognitive Radial Basis Function Neural Classifier (EMcRBFN). An automatic whole brain Voxel Based Morphometry (VBM) approach is used to identify gray matter composition in the brain from structural Magnetic Resonance Imaging (MRI) and an improved *q*-Gaussian classifier and its metacognitive learning algorithm has been proposed to approximate the functional relationship between the high dimensional VBM features and the true class labels. Recent genetic studies indicate that ASD manifests in different ways between males and females and also between adolescents and adults. Accordingly, the proposed EMcRBFN classifier has been evaluated using the publicly available Autism Brain Imaging Data Exchange dataset with a comprehensive study on both males and females and also between adolescents and adults in both categories. EMcRBFN classifier performance is compared with currently existing results for ASD classification in the literature and also with well known standard classifiers. The results clearly indicate that the performance of the EMcRBFN classifier is better than that of the other classifiers considered in this study. Further, the comprehensive study also indicates that the following subregions in the brain viz., premotor cortex and supplementary motor cortex are affected for adult-females while the somatosensory cortex subregion is affected for adolescent-females with ASD. Similar results indicate that the precentral gyrus, motor cortex, medial frontal gyrus and the paracentral lobule areas are affected for adolescent males while the superior frontal gyrus and the frontal eye fields areas are affected for adult males with ASD.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Autism Spectrum Disorder (ASD), is a highly genetic neurodevelopmental condition that affects nearly 1 in 68 children (CDC, 2014). ASD is characterized by impaired social communication, social reciprocity and repetitive stereotyped behavior. Motor function, attention and other cognitive domains may also be affected (APA et al., 2013; Eisenmajer et al., 1996; Volkmar, Cohen, Bregman, Hooks, & Stevenson, 1989). The main drawback of traditional diagnosis methods such as Autism Diagnostic Observation Schedule (Lord, 1989) and Autism Diagnostic Interview (Lord, Rutter, & Le Couteur, 1994) is that they are heavily dependent on the skill of the examiner in eliciting a developmental history accurately from informants who know the

subject well. Since ASD is a developmental disorder, the symptoms observed at the time of the diagnosis may have been suppressed due to the coping strategies developed by the subject over time. In such cases, retrospective information may be needed, which may not always be available. Since these methods depend heavily on the availability of an informant and the accuracy of the information provided, they are error-prone. Another problem with these methods is their inability to point out the biological etiology (eg. genetic, environmental, etc.) that is responsible for a particular ASD related condition in a subject, since many etiological factors can lead to the same behavioral phenotype (Geschwind & Levitt, 2007). Hence, there is a need for a method of ASD detection that is non-invasive, automatic and is also capable of pointing out some biological factors responsible for different behavioral conditions. One such method is by using brain imaging as it can provide insights into the underlying brain structure and composition. Magnetic Resonance Imaging (MRI) is the most important brain imaging procedure that provides accurate information about the shape and volume of the brain.

* Corresponding author. Tel.: +65 6790 6185; fax: +65 6792 6559.

E-mail addresses: vigneshw1@e.ntu.edu.sg (V. Subbaraju), ssundaram@ntu.edu.sg, suresh99@gmail.com (S. Sundaram), ensundara@ntu.edu.sg (S. Narasimhan), bsmahanand@gmail.com (M.B. Suresh).

When analyzing the brain using MRI, anatomical features of the brain such as regional and volumetric compositions of gray matter and white matter can be obtained using structural MRI while certain other aspects such as connectivity can be obtained using functional MRI. Thus, structural and functional MRI are two complementary and vital methods of analyzing brain images. Among the different components of the brain, gray matter is chiefly responsible for cognitive functions and it is known that ASD patients suffer from deficiencies in social cognition. Hence, there is a need to study the differences in gray matter composition obtained via structural MRI for effective classification of ASD patients and healthy controls.

Previous studies using structural MRI of ASD patients have observed regional differences in gray matter composition in the areas around the frontal cortex like middle frontal gyrus, inferior frontal gyrus and medial orbito frontal cortex (Bonilha et al., 2008; Hadjikhani, Joseph, Snyder, & Tager-Flusberg, 2006; Hardan et al., 2006; Hardan, Libove, Keshavan, Melhem, & Minshew, 2009; Hyde, Samson, Evans, & Mottron, 2010; McAlonan, et al., 2005; Waiter et al., 2004) and volumetric differences in gray matter around the superior temporal sulcus, inferior parietal lobule, cingulate and fusiform gyrus (Boddaert et al., 2004; Bonilha et al., 2008; Hadjikhani et al., 2006; McAlonan et al., 2005; Rojas et al., 2006; Waiter et al., 2004). These studies explored gray matter differences in specific regions of the brain, however, in the case of MRIs from children whose brain is still developing fast, the specific regions may not be fully grown to facilitate a clear demarcation of their boundaries. Hence, whole brain analysis is preferred for the ASD detection problem since it gives the ability to explore the entire brain for abnormalities. Whole brain MRI based analysis for ASD detection has been studied in recent works such as (Calderoni et al., 2012; Ecker et al., 2010b) and both these methods have attempted to use voxel-wise comparison of the brain composition among the ASD patient and healthy controls. Voxel-Based Morphometry (VBM) is an automatic image analysis approach for identifying the differences in the amount of gray matter or white matter between the normal subjects and ASD patients. In analyzing the features from VBM, multi-variate methods are preferred against mass-univariate methods (Ecker et al., 2010a) making a strong case for machine learning based approaches. The extracted features from VBM are directly used to build classifiers. In (Ecker et al., 2010b), both gray matter and white matter composition of 44 male adults were studied and it was found that gray matter offers better accuracy (81%) than white matter (68%). In (Calderoni et al., 2012) only gray matter was used to classify ASD in 76 female children with an area under the curve (AUC) of 0.80. However, these studies are based on proprietary data with very limited number of samples.

Recently, the Autism Brain Imaging Data Exchange (ABIDE, 2013) consortium was formed and a large dataset consisting of MRI volumes obtained from various locations was made publicly available, providing an unprecedented opportunity for large scale investigation of ASD from MRI. Most of the studies based on this open dataset (Idiaka, 2015; Nielsen et al., 2013) have focused on using functional MRI to investigate the connectivity patterns of different brain regions. In (Nielsen et al., 2013) the entire brain was divided into 7266 sub-regions and the temporal correlation of the blood oxygen level dependent (BOLD) signal between the all possible pairs of regions was used as a measure the connectivity. This resulted in a huge number of 24 million features and a low classification accuracy 60% has been reported. In (Idiaka, 2015) only a subset of the ABIDE dataset is investigated for functional connectivity of the regions of the brain as specified by the Automatic Anatomical Labeling (AAL) template. While a high accuracy of 90% has been reported, an accurate demarcation of the AAL regions might not be possible in all cases. Functional MRI can provide insights into certain aspects of brain function such as connectivity, but they are unable to point out anatomical abnormalities

in the brain due to ASD and the recent studies using ABIDE dataset have not been able to directly link the differences in functional connectivity to the observed behavioral symptoms. In (Vigneshwaran, Mahanand, Suresh, & Savitha, 2013), a limited part of the ABIDE dataset, contributed by New York University was used to study ASD classification performance when using gray matter based features that were extracted from VBM. An accuracy of 78% was achieved indicating the feasibility of the approach.

Recent genetic studies (Jacquemont et al., 2014) suggest a 'female protective model' in the manifestation of ASD and it is estimated that the incidence of ASD is at least five times higher in males than in females (APA et al., 2013). This is motivating researchers to analyze females and males separately and some VBM studies have suggested that females and males may constitute different autism groups (Lai et al., 2013). Since ASD is much more common in males, most of the studies have focused on male subjects and there is very limited literature available for females as noted in a recent study on females with ASD (Calderoni et al., 2012). Similarly, since ASD is known to be a developmental disorder and longitudinal studies have found age-specific anatomic anomalies (Courchesne, Campbell, & Solso, 2011a; Courchesne, Webb, & Schumann, 2011b), it is necessary to consider adults and adolescents separately. Hence, there is a need for separate age and gender based studies on the structural MRI available in the complete ABIDE dataset in order identify differences in gray matter (which may be responsible for impaired social cognition in ASD) and use it for effective classification of ASD patients.

Earlier, ASD classification studies have been conducted using VBM detected features and Support Vector Machine (SVM) classifier (Calderoni et al., 2012; Ecker et al., 2010b; Orru, Pettersson-Yeo, Marquand, Sartori, & Mechelli, 2012). Recently a 'Projection Based Learning Metacognitive Radial Basis Function Network' (PBL-McRBFN) was proposed (Sateesh Babu & Suresh, 2013) for Parkinson's disease detection effectively and have been shown to perform better than SVM. PBL-McRBFN uses the concept of meta-cognition from learning psychology area to select the samples based on the information content in them instead of using all the samples. In (Vigneshwaran et al., 2013; Vigneshwaran, Suresh, Mahanand, & Sundararajan, 2015a), the performance of PBL-McRBFN for ASD classification was evaluated and the results indicated better performance than conventional classification algorithms such as SVM, Naive Bayes classifier, decision trees and nearest neighbor classifier. However, PBL-McRBFN uses a Gaussian activation function, whose response drops sharply in the tail region, when the feature dimension is high. This phenomenon results in a high number of neurons being added to the network and affects its generalization performance. This phenomenon is also known to cause some performance issues in certain other application areas such as real-time signal processing and image retrieval (Fernández-Navarro, Hervás-Martínez, Gutiérrez, Peña-Barragán, & López-Granados, 2012). Hence, it is desirable to use a more general radial basis function such as the q -Gaussian activation function which can assume the form of a Gaussian, Cauchy or inverse-multiquadratic function according to the parameter q . The q -Gaussian activation function is known to be effective for classification problems (Babu, Suresh, & Mahanand, 2013; Savitha, Suresh, & Kim, 2014; Vigneshwaran, Suresh, & Sundararajan, 2015b).

In order to address the above mentioned issues in the ASD classification literature, in this paper, detailed age and gender specific studies have been performed on the structural MRI data available on the complete ABIDE dataset to detect differences in gray matter composition and use it for accurate classification. VBM has been used to extract features based on gray matter composition and the brain regions affected for different age and gender groups are highlighted. The q -Gaussian activation function is used to improve the performance of the PBL-McRBFN and the resultant algorithm, called as the Extended-Metacognitive Radial Basis Function Network (EMcRBFN), is used for classification with improved accuracy.

VBM analysis of MRI from all the females of the ABIDE dataset show difference in the gray matter, composition, in the paracentral lobule and the motor cortex regions of the brain. It was also found that EMcRBFN is able to achieve about 3% higher accuracy than the PBL-McRBFN and about 11% higher accuracy than that of SVM. Among females, it is further shown that different features are obtained for adults and adolescents when performing VBM analysis and hence different regions of the brain needs to be considered for these two groups. In adults, activations are found in the premotor cortex and supplementary motor cortex. Premotor cortex plays a role in planning, controlling, and executing voluntary movements and the supplementary motor cortex plays a role in execution of complex movements. For adolescents, activations are found in the somatosensory cortex, which is responsible for touch sensations. These regions seem to be consistent with literature and the symptoms observed in ASD patients. It is also shown that the ASD classification performance can be improved by 13% for adults and 6% for adolescents by performing separate feature extraction and building separate classification models.

Similarly, when considering all the males in the ABIDE dataset differences in the gray matter composition are found in the caudate, premotor cortex, supplementary motor cortex and the medial frontal gyrus regions of the brain. It was also found that the PBL-McRBFN is able to achieve about 3% higher accuracy than the popular SVM. It is then shown that the overall ASD classification accuracy can be improved by about 10% by performing separate feature extraction and building separate classification model for adults. But the same strategy does not significantly improve the classification performance for male adolescents due to the high variability in the gray matter composition of male adolescent brain which is still developing fast following a unique developmental path. Whole brain VBM analysis indicates that the precentral gyrus, motor cortex, medial frontal gyrus and the paracentral lobule areas are the most affected regions for adolescent males while the superior frontal gyrus and the frontal eye fields are the regions affected for adult males. The cognitive functions of these areas seem to match some of the symptoms observed in ASD patients.

This paper is organized as follows: [Section 2](#) briefly describes the dataset used, the framework of the proposed method, the feature extraction process using VBM and the working principles of the EMcRBFN classifier. [Section 3](#) describes the main results from this study using the EMcRBFN classifier and a comparison with other classification methods such as SVM, PBL-McRBFN, ELM and Multi-Layer Perceptron. [Section 4](#) summarizes the conclusion from this study.

2. Materials and methods

2.1. MRI acquisition

MRI scans used in this study were obtained from the openly available, Autism Brain Imaging Data Exchange (ABIDE) data set. It consists of MRI scans contributed from a total of 16 international sites, yielding 1112 datasets composed of both MRI data and an extensive array of phenotypic information common across nearly all sites. Among the 1112 samples, only 1054 samples were used in this study and the rest were rejected due to improper segmentation. Since, recent genetic studies ([Jacquemont et al., 2014](#)) have suggested separate gender-specific approaches for the screening of neurodevelopmental disorders, the 154 MR images that belong to females and the 900 males in this dataset are considered in separate studies in this paper. The details of the data used in this study are summarized in [Table 1](#).

A high percentage of the samples are from children and adolescents below the age of 18. Since ASD is a developmental disorder, age can be an important factor affecting the diagnosis. Hence in this study separate VBM analysis is performed for adults and adolescents and separate classifiers are built for these age groups.

Table 1
Details of data used in the study.

	Female		Male	
	Healthy persons	ASD patients	Healthy persons	ASD patients
Adults- (>18yrs)	23	15	159	141
Adolescents- (<18yrs)	72	44	292	308
Total	95	59	451	449

2.2. Feature extraction using VBM

Voxel based morphometry ([Ashburner & Friston, 2000](#)) is an automatic image analysis approach for identifying the amount of gray matter or white matter differences between ASD patients and normal persons. The steps involved in VBM to identify significant differences between the groups of persons are: spatial normalization, segmentation, smoothing and statistical analysis as shown in [Fig. 1](#). In the unified segmentation step, tissue segmentation, bias correction and image registration are combined in a single general model ([Ashburner & Friston, 2005](#)). The VBM approach is used in this paper to identify the regional differences in gray matter between groups of persons and to extract morphometric features from MR images. The segmented and registered gray matter images are then smoothed by convolving with a 10 mm full-width at half-maximum isotropic Gaussian kernel. For statistical testing, the whole brain group comparison between normal persons and ASD patient is conducted with a two-sample t-test and gray matter volume is used as the covariate.

The output of VBM analysis is the maximum intensity projection, which highlights the regions that show significant difference in gray matter between ASD patients and normal persons. In order to obtain the feature vector corresponding to a particular MRI scan, the voxel location of significant regions in the MIP is used as a mask to extract the features (gray matter tissue probability values). The extracted features from all the segmented gray matter images are arranged linearly in lexical order and then used as an input to the EMcRBFN classifier.

2.3. EMcRBFN classifier

ASD detection problem can be viewed as a mapping from the domain of the features to the domain of the class labels. A classifier is needed to approximate the underlying decision function that maps $\mathbf{x}^t \in \mathbb{R}^m \rightarrow \mathbf{y}^t \in \mathbb{R}^n$. Recently, the Metacognitive Radial Basis Function Network Classifier (McRBFN) and its Projection Based Learning (PBL) algorithm was proposed in ([Sateesh Babu & Suresh, 2013](#)) and successfully applied for Parkinson's disease detection. PBL-McRBFN is a neural network with a single hidden layer of neurons with radial basis activation function. It begins with zero hidden neurons and selects a suitable strategy for each sample (adds neurons or updates output parameters or deletes sample or reserves sample) to achieve its objective. PBL-McRBFN uses the Gaussian activation function in the hidden layer. While the Gaussian RBF function is very common, its response is known fall sharply in the tail region in high-dimensional feature space. This leads to the additions of more neurons in the network and poor generalization performance. Hence, certain other radial basis functions are preferred in some applications. For example, the Cauchy radial basis function is preferred in image retrieval and the inverse multi-quadratic radial basis function is preferred for some real-time signal processing applications ([Fernández-Navarro et al., 2012](#)). In order to accommodate such variations, it is desirable to employ a radial basis function with kernels, whose shape can be relaxed or contracted. In this work, it is proposed to modify the McRBFN by using a q -Gaussian activation function instead of a Gaussian activation function along with an improved learning algorithm and apply it

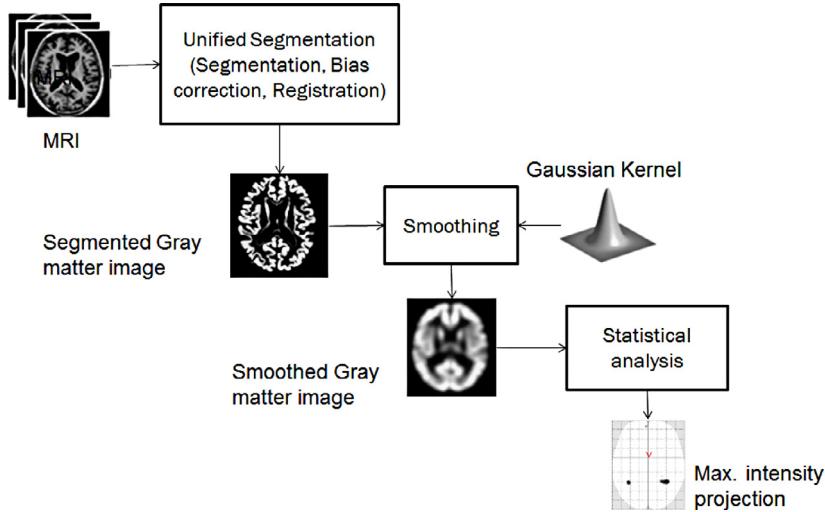


Fig. 1. Voxel based morphometry- steps.

to the ASD detection problem and resultant network is named as the Extended Metacognitive Radial Basis Function Network (EMcRBFN).

2.3.1. Architecture

The EMcRBFN is a neural network with a single hidden layer whose neurons have a radial basis activation function. The most common radial basis function is the Gaussian activation function which was used in PBL-McRBFN. One of the characteristics of the Gaussian radial basis function is that it provides a very small response when the input is far away from the center and falls in the tail region of the Gaussian activation function. This effectively renders a neuron to become inactive in some regions of the feature space. While this is a desirable attribute in some applications, in some other applications this might be undesirable. This affects the generalization performance of the RBF network and this issue has been discussed in detail in (Fernández-Navarro et al., 2012). This issue can be addressed by employing an activation function such as the q -Gaussian function, whose kernels can be parameterized to adjust the shape of the tail as desired.

The response of the neurons in the hidden layer employing a q -Gaussian RBF is given by:

$$h_k^t = \begin{cases} (r_k^t)^{\frac{1}{1-q}}, & \text{if } r_k^t > 0 \\ 0 & \text{otherwise} \end{cases}; \quad k = 1, \dots, K \quad (1)$$

$$\text{where, } r_k^t = (1 - (1 - q) d_k(\mathbf{x}^t)) \quad \text{and} \quad d_k(\mathbf{x}^t) = \frac{\|\mathbf{x}^t - \boldsymbol{\mu}_k^l\|^2}{(\sigma_k^l)^2} \quad (2)$$

d_k is the radial distance of the sample from the k th neuron, $\boldsymbol{\mu}_k^l \in \mathbb{R}^m$ is the center and $\sigma_k^l \in \mathbb{R}^+$ is the width of the k th hidden neuron. Here, the superscript l represents the class of that hidden neuron to which it belongs. The parameter $q \in \mathbb{R}^1$ is used to produce the various RBF functions. For example, when $q \rightarrow 1$ the q -Gaussian converges to a Gaussian RBF, while $q \rightarrow 2$ and $q \rightarrow 3$ converge to the Cauchy RBF and the inverted multi-quadratic RBF, respectively. The response of the q -Gaussian activation function in one-dimensional space is shown in Fig. 2 for various values of q . It can be observed that the tail of the Gaussian can be manipulated by choosing appropriate values of q . As q increases, the activation for samples far from the center of the RBF is greater than those of the Gaussian RBF for the same samples. In practical applications, the q parameter is chosen in the range of [0,3].

The EMcRBFN starts with zero hidden neurons and adds neurons whenever necessary and thus automatically determines the number

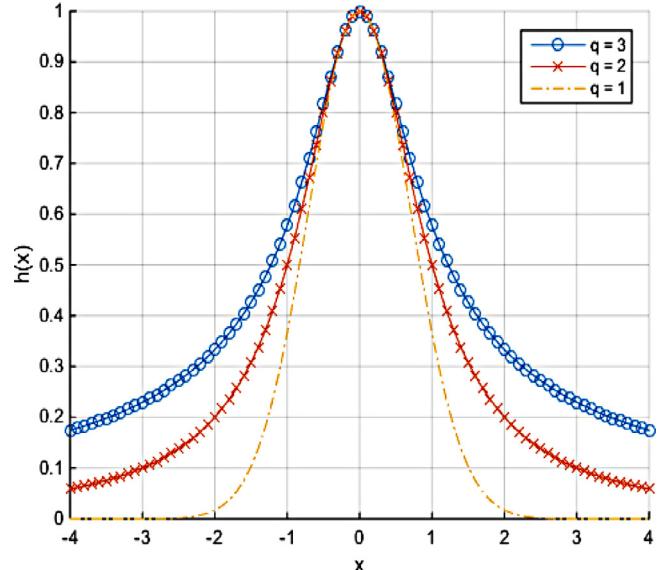


Fig. 2. Responses of q -Gaussian function in 1-D space for different values of q ($\mu = 0$; $\sigma = 1$).

of hidden neurons. We can assume, without loss of generality, that the EMcRBFN builds K q -Gaussian neurons from $t - 1$ training samples. For a given input \mathbf{x}^t , the predicted output \hat{y}_j^t of EMcRBFN is

$$\hat{y}_j^t = \sum_{k=1}^K w_{kj} h_k^t, \quad j = 1, 2, \dots, n \quad (3)$$

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 the Eqs. (1) and (2).

2.3.2. Sequential Learning algorithm of EMcRBFN classifier

The parameters of the EMcRBFN are chosen by a sequential learning process that is an extended version of the PBL algorithm proposed in (Sateesh Babu & Suresh, 2013). The following are the two modifications to the PBL algorithm proposed in this paper.

- Whenever a neuron is added to the network, the q parameter of the neuron has to be initialized.

- When adding a neuron or updating the parameters of the network, an additional check is performed whether any neuron was fired with sufficiently strong response.

Other than these two modifications, the working principles and the mathematical derivations are exactly the same as for the PBL learning algorithm and readers can refer to (Sateesh Babu & Suresh, 2013) for detailed derivations. A summary of the working principles of the learning algorithm is given below and the modifications proposed in this paper are highlighted.

2.3.2.1. Overview of the PBL algorithm. The PBL algorithm formulates the sum of squared errors (SSE) at the McRBFN output neurons as its energy function and finds the network output parameters for which the energy function is minimum. The SSE at the McRBFN output layer is given by,

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

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

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

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

In order to obtain the optimal \mathbf{W}^* corresponding to the minimum energy point of the energy function ($J(\mathbf{W}^*)$) the first order partial derivative of $J(\mathbf{W})$ with respect to the output weight is equated to zero, and as derived in (Sateesh Babu & Suresh, 2013). This results in the solution,

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

where the projection matrix $\mathbf{A} \in \mathbb{R}^{K \times K}$ is given by

$$a_{kp} = \sum_{i=1}^t h_k^i h_p^i, k = 1, \dots, K; p = 1, \dots, K \quad (7)$$

and the output matrix $\mathbf{B} \in \mathbb{R}^{K \times n}$ is

$$b_{pj} = \sum_{i=1}^t h_k^i y_j^i, p = 1, \dots, K; j = 1, \dots, n \quad (8)$$

The PBL-McRBFN in (Sateesh Babu & Suresh, 2013) uses the following measures of knowledge for every training sample, namely,

Estimated Class label (\hat{c}^t):

$$\hat{c}^t = \arg \max_{j \in 1, 2, \dots, n} \hat{y}_j^t \quad (9)$$

Maximum Hinge Error (E^t):

$$E^t = \max_{j \in 1, 2, \dots, n} |e_j^t| \quad (10)$$

Where,

$$e_j^t = \begin{cases} 0 & \text{if } y_j^t \hat{y}_j^t > 1 \\ y_j^t - \hat{y}_j^t & \text{otherwise} \end{cases} \quad j = 1, 2, \dots, n \quad (11)$$

Confidence of Classifier ($\hat{p}(c^t | \mathbf{x}^t)$):

$$\hat{p}(j | \mathbf{x}^t) = \frac{\min(1, \max(-1, \hat{y}_j^t)) + 1}{2}, j = c^t \quad (12)$$

Class-wise Significance (ψ_c):

$$\psi_c = \frac{1}{K^c} \sum_{k=1}^{K^c} h(\mathbf{x}^t, \boldsymbol{\mu}_k^c) \quad (13)$$

Where, K^c be the number of neurons associated with the class c . For more details on the class-wise significance of McRBFN, one can refer to (Sateesh Babu & Suresh, 2012). Using these measures, a suitable

strategy is selected for each training sample when it is presented to the network for learning. The four learning strategies are, viz. sample deletion, neuron addition, parameter update and sample reserve strategy.

2.3.2.2. Extensions to the PBL algorithm. In this paper the PBL algorithm has been extended for the following two reasons. Firstly, the EMcRBFN uses a q -Gaussian activation function and it is necessary to initialize the q -parameter for each neuron. Secondly, the PBL algorithm does not add any neuron as long as the class is estimated correctly and the maximum hinge error is low. This is not sufficient to determine whether the network effectively represents the knowledge in a sample. There can be a situation where the input is far away from the centers of all the neurons in the network and hence none of the neurons in the fire with sufficient response. However, since the hinge error is used to determine the class label, it might be possible that the network predicted the class label accurately by chance. Thus, in this case even though the class label was predicted accurately, the confidence of the classifier will be very low. In order to catch such conditions, the EMcRBFN uses the maximum firing strength ϕ_{max} as an additional knowledge measure for selection of learning strategies.

Maximum Firing Strength (ϕ_{max}): for each input, the maximum firing strength ϕ_{max} is the highest response obtained from any of the hidden neurons in the network. If the value of ϕ_{max} is too low for a particular sample, then it means that the knowledge in this sample is not sufficiently represented by the network.

This new knowledge measure is used in the criteria for neuron addition and parameter update strategies as explained below.

Neuron Growth Strategy: a new neuron should be added when a training sample contains significant information and the knowledge contained in the sample is not adequately represented by the existing network. The neuron growth criterion is given by

$$(\hat{c}^t \neq c^t \text{ OR } E^t \geq \beta_a \text{ OR } \phi_{max} \leq \beta_f) \text{ AND } (\psi_c(\mathbf{x}^t) \leq \beta_c) \quad (14)$$

where β_c is the meta-cognitive knowledge measurement threshold and β_a is the self-adaptive meta-cognitive addition threshold as described in (Sateesh Babu & Suresh, 2013). β_a is adapted in each training step according to

$$\beta_a := \delta \beta_a + (1 - \delta) E^t \quad (15)$$

where δ is the slope that controls rate of self-adaptation and is set close to 1.

In addition to these two thresholds, a firing threshold (β_f) is introduced in the EMcRBFN. At each training step, the EMcRBFN checks if any of the existing neurons provide sufficient response above this firing threshold. Even if the network may be able to correctly classify the sample, if none of the neurons have sufficient firing strength, then it might be necessary to add a neuron for this sample in order to represent the knowledge in the sample effectively and thus improve the confidence of the classifier.

Initialization of q -parameter: the value of the parameter q for each neuron is initialized when a hidden neuron $K + 1$ is added to the network. This initialization done according to the following logic. If there are currently no hidden neurons in the network i.e the training process has just begun, the parameter q for the new hidden neuron is chosen from a uniform distribution in the range of (0.5, 1.4]. However if there are already some neurons present in the network, three random values are picked from uniform distributions in the ranges of (0.5, 1.4], (1.4, 2.2] and (2.2, 3.0]. Among these three random values, the value that results in the least RMS error for the network is chosen as the parameter q for the new neuron. Each neuron is thus assigned a particular value of q when it is added to the network. The different ranges from which the parameter q is picked ensure that there is variety in the shape of the activation functions in the neural network. Breaking down the ranges further might result in better initialization of the q -parameter at the cost of additional computation.

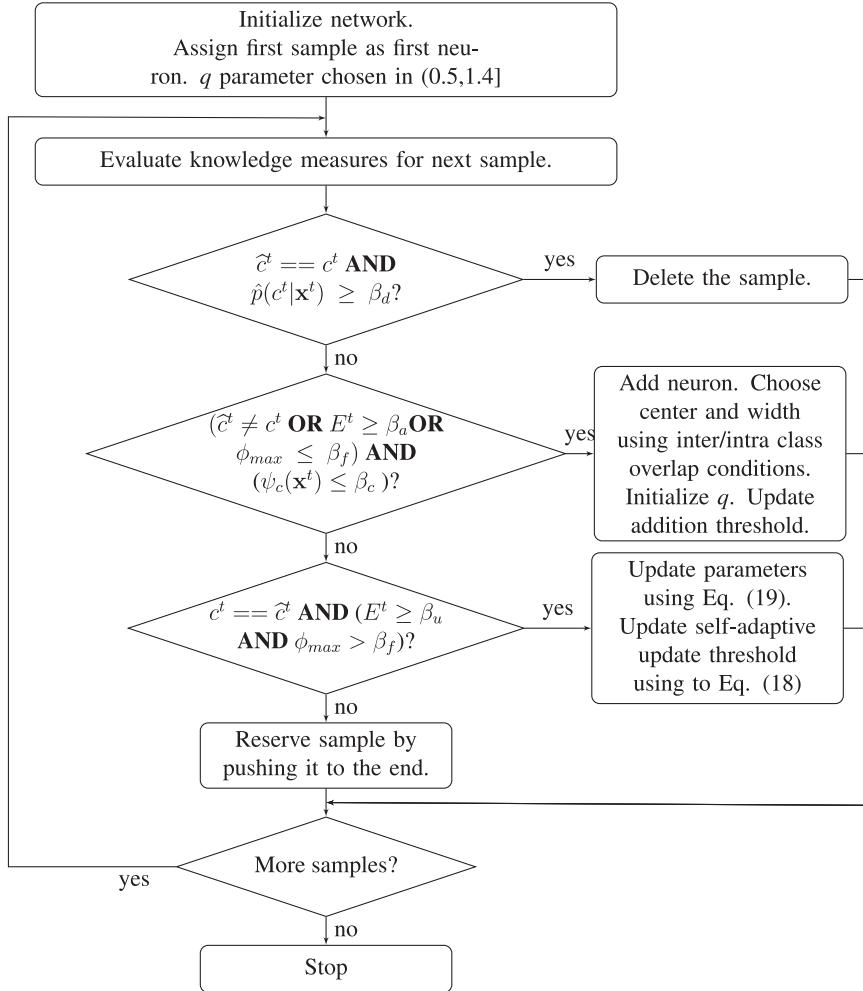


Fig. 3. Summary of the learning process of the EMcRBFN algorithm.

Initialization of center and width: when a new hidden neuron $K + 1$ is added, the center and width parameters are initialized using the overlapping and distinct cluster criterion. The new training sample may either be covered partially by existing neurons (overlap) for other classes or it could be from a distinct cluster far away from the nearest neuron in the same class. These conditions affect the classification performance of a classifier significantly. Hence, the EMcRBFN measures inter/intra class nearest neuron distances from the current sample in assigning the new neuron parameters as done in the PBL (Sateesh Babu & Suresh, 2013).

Re-Estimation of weights: the output weights are re-estimated whenever a new hidden neuron is added as done in (Sateesh Babu & Suresh, 2013). The equation to estimate the output weights is reproduced here,

$$\begin{bmatrix} \mathbf{W}_K \\ \mathbf{w}_{K+1} \end{bmatrix} = (\mathbf{A}_{(K+1) \times (K+1)})^{-1} \mathbf{B}_{(K+1) \times n} \quad (16)$$

where the matrices in the right hand side can be obtained from Eqs. (7) and (8). One can then use matrix identities to avoid re-computation of the inverse as shown in (Babu & Suresh, 2013).

Parameters update strategy: the current (t th) training sample is used to update the output weights ($\mathbf{W}_K = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_K]^T$) of the EMcRBFN, if the following criterion is satisfied.

$$c^t == \hat{c}^t \text{ AND } E^t \geq \beta_u \text{ AND } \phi_{\max} > \beta_f \quad (17)$$

where β_u is the self-adaptive meta-cognitive parameter update threshold. The initial value of β_u can be selected in the interval [0.4 -

0.7]. The value of β_u is adapted based on the prediction error as:

$$\beta_u := \delta \beta_u + (1 - \delta) E^t \quad (18)$$

The output weights are updated as follows:

$$\mathbf{W}_K = \mathbf{W}_K + \mathbf{A}^{-1} (\mathbf{h}^t)^T (\mathbf{e}^t)^T \quad (19)$$

The training process aims to find the minimum value of the objective function that is based on the hinge loss error. The introduction of a q -Gaussian activation function does not affect the method of calculation of the weights that minimize the objective function.

A summary of the EMcRBFN algorithm is given in the form of a flowchart in Fig. 3.

3. Experimental results

In this section, the problem of detecting ASD from gray matter composition is studied experimentally using the publicly available ABIDE (ABIDE, 2013) dataset. Separate studies for males and females are presented and for each gender, separate studies are presented for the detection of ASD among adolescents (age < 18 years) and adults (age ≥ 18 years). Based on these studies the brain regions that are affected for these age groups have been identified. Finally, classification performance for ASD detection using the EMcRBFN classifier is compared with PBL-McRBFN and also the well known SVM, MLP and ELM classifiers.

3.1. Dataset and algorithms

The ABIDE dataset consists of MRI volumes from 1112 subjects. Out of these only 1054 of them were considered in this paper while the others were rejected due to improper segmentation during VBM analysis. Within these 1054 subjects, there were 154 female subjects, of which 95 are normal persons and 59 are ASD patients. There were 900 male subjects, of which 451 are normal persons and 449 are ASD patients. VBM analysis is used to extract features from these MRI scans and various algorithms are used to build classifiers from the extracted features and their performance is compared. VBM analysis was performed using the Statistical Parametric Map (SPM) software package (SPM8, 2011). The output of VBM analysis is the maximum intensity projection (MIP), which is a map that highlights the voxels that show significant difference in the composition of the brain. These VBM detected brain regions are used as masks for feature extraction. In order to study the impact of age and gender, six separate VBM analysis and classification experiments were performed:

- Case A: MRI from all females put together.
- Case B: considering MRI from adolescent females ($\text{age} < 18$) only.
- Case C: considering MRI from adult females ($\text{age} \geq 18$) only.
- Case D: MRI from all males put together.
- Case E: considering MRI from adolescent males ($\text{age} < 18$) only.
- Case F: considering MRI from adult males ($\text{age} \geq 18$) only.

EMcRBFN proposed in the previous section is used to build a classification model based on the feature vectors that were extracted from VBM analysis. For comparison of the classification performance, the following algorithms have been used in this study;

- Support Vector Machine (SVM). This is the most common classifier used in literature. In this study, a Gaussian kernel is chosen for the SVM and the parameters of the SVM were optimized using the gridsearch method. The LIBSVM (Chang & Lin, 2011) is a popular implementation of the SVM with gridsearch method (Hsu, Chang) and it has been used in the simulation studies reported in this paper.
- Multi Layer Perceptron (MLP). This is a simple feed-forward neural network where the neurons in the hidden layers have a sigmoid activation function and the neurons in the output layer have a softmax activation function. The network is trained using the scaled-conjugate gradient back propagation algorithm. This can be implemented using the ‘*patternnet*’ function in MATLAB.
- Extreme Learning Machine (ELM). This is a recently popular (Huang, Zhu, & Siew, 2006) neural network algorithm with a single hidden layer, which drastically reduces training time by skipping iterative tuning steps and using an analytical method of determination of weights. In this study, RBF neurons are used in the hidden layer and the publicly available MATLAB implementation of the ELM is used for our simulations.
- PBL-McRBFN classifier (Sateesh Babu & Suresh, 2013). This is a RBF neural network based approach employing meta-cognitive learning principles.

For this study, in each trial, 75% of the samples are chosen randomly as the training set and the rest as testing set. Performance evaluation of the classifier has been done by generating 10 such random trials. All the algorithms were implemented in MATLAB and the simulations were run on a laptop with Intel Core i7 quad core processor, 2.1 GHz CPU and 16 GB RAM.

3.2. VBM analysis for ASD detection in females

In this section, the results of the VBM analysis for females i.e the three Cases A, B and C mentioned in Section 3.1 are provided and discussed. For each case, the maximum intensity projection, the lo-

Table 2
VBM detected regions in females for different age-groups.

Age-group	Identified regions	Functionality
Adolescent	Somatosensory cortex	Touch
Adult	Premotor cortex	Planning, controlling, and executing voluntary movements
All	Supplementary motor cortex	Executing complex movements
	Primary motor cortex	Execution of movement
	Paracentral lobule	Motor and sensory stimulations

cation of significant voxels and the number of features obtained are given (Table 2).

3.2.1. Case A: considering all females

As the first experiment in VBM analysis, all the MRI scans belonging to females in the ABIDE dataset are taken and a VBM analysis is done on this whole set. The aim of this study is to identify the regions of the brain that show significant differences in the gray matter composition. From these VBM identified regions, the gray matter probabilities are taken as the features. These features are then used by the different classifiers for ASD detection.

Using the VBM analysis, maximum intensity projections ($P < 0.001$, uncorrected) of significant areas that show significant differences in the gray matter density between normal persons and ASD patients are obtained. In order to locate the regions obtained from maximum intensity projection with respect to the spatial locations in the brain, a three-dimensional volume-rendering is performed. The rendered image shows the areas with significant difference in gray matter density between normal persons and ASD patients. These results are presented in Fig. 4. Fig. 4a–c present the sagittal, coronal and axial views of the brain respectively. In Fig. 4a, subfigure (i) shows the MIP, where the significant regions are highlighted in black. The corresponding three-dimensional volume-rendered display is shown in subfigures (ii) and (iii) which are the views from the right side and the left side of the brain respectively. In the subfigures (ii) and (iii), the affected areas are highlighted in red. Similarly, in Fig. 4b, the subfigure (i) represents the MIP and the subfigures (ii) and (iii) represent three-dimensional volume-rendered display as viewed from the front and back of the brain respectively. In Fig. 4c, subfigures (ii) and (iii) represent the three-dimensional volume-rendered display in axial view from the bottom and the top of the brain, respectively.

When MRI scans of all the females in the dataset are considered as in Fig. 4, it may be noted that the paracentral lobule, primary motor cortex regions of the brain, show significant difference in gray matter composition for ASD patients. From these voxel locations, the gray matter tissue probabilities are taken as the features resulting in 66 features for each MRI scan.

3.2.2. Case B: ASD detection for adolescent females ($\text{age} < 18$ years)

Since ASD is a developmental disorder, it may affect people at different ages in a different manner. Hence, in this section, all the MRI scans belonging to adolescent females (< 18 years) in the ABIDE dataset are considered separately for VBM analysis. The regions of the brain that show significant differences in the gray matter composition for adolescent females are identified by VBM.

Similar to Section 3.2.1, the maximum intensity projections ($P < 0.001$, uncorrected) of significant areas with changes in gray matter density between normal persons and ASD patients are obtained from the VBM analysis for adolescent females and shown in Fig. 5. From Fig. 5, it can be seen that the somatosensory cortex region is affected for adolescent females. This region is part of the larger motor cortex area and it is involved in touch sensation. By considering the

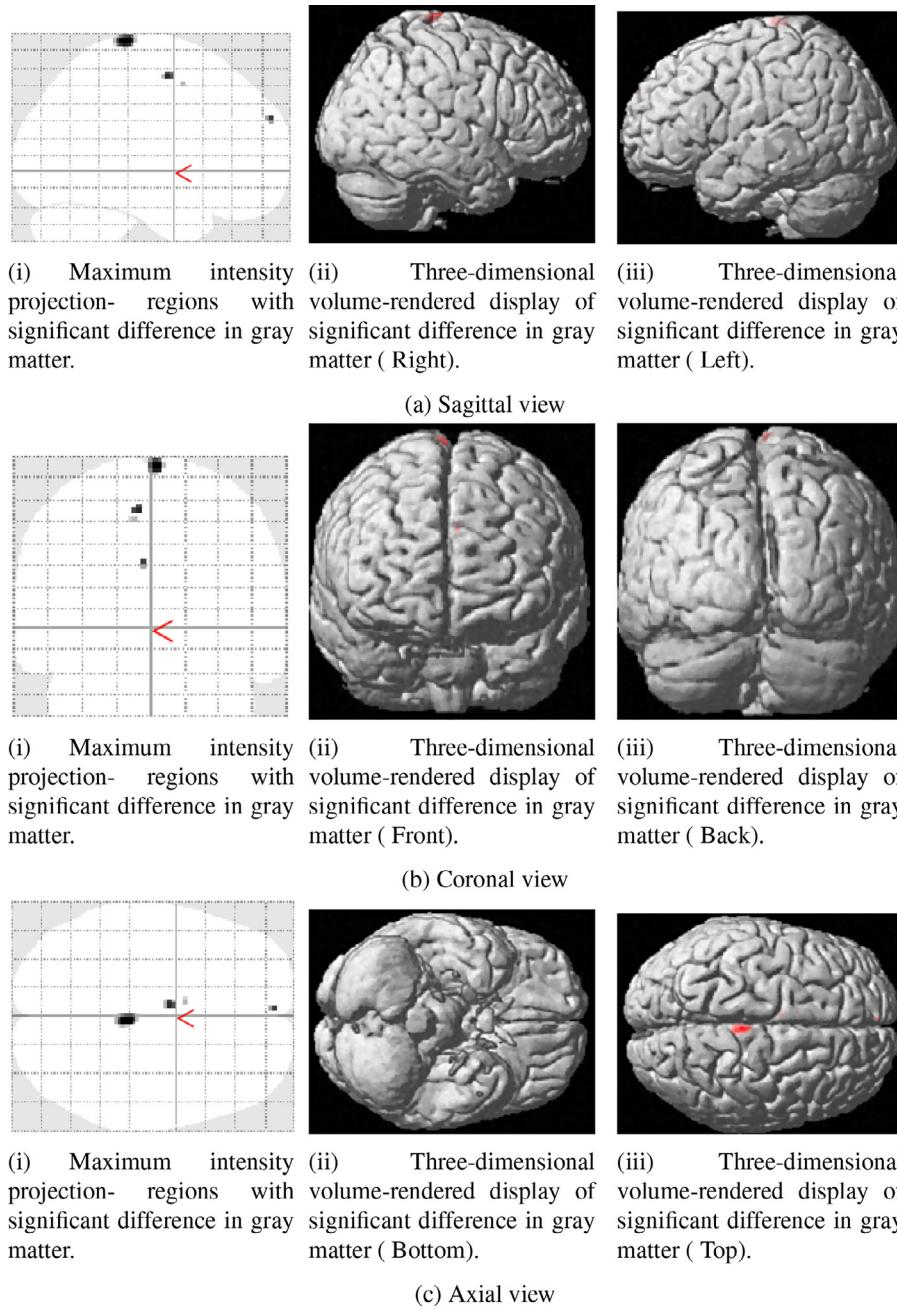


Fig. 4. Brain regions with significant differences in gray matter density in ASD patients, when considering MRI scans of all females in the dataset.

gray matter tissue probability, the analysis resulted in 104 features for each MRI scan.

3.2.3. Case C: ASD detection for adult females ($age >= 18$ years)

In this section, the MRI scans from the female adults ($Age >= 18$ years) in the ABIDE dataset are considered separately. The condition of patients with ASD are known to improve with age even though they may not fully recover. Hence, it is possible that different regions of the brain are affected for adult females and adolescent females. VBM analysis is performed as done in the Sections 3.2.1 and 3.2.2 in order to identify the brain regions that show significant differences in the gray matter composition for adult females suffering from ASD. The maximum intensity projections ($P < 0.001$, uncorrected) of significant areas are shown in Fig. 6. From Fig. 6, it can be seen that the premotor cortex and supplementary motor cortex areas are affected for adult females. Premotor cortex plays a role in direct control of

behavior and also in planning, controlling, and executing voluntary movements. Supplementary motor cortex plays a role in execution of complex movements. 43 features for each MRI scan are obtained by considering the gray matter probability in these areas.

3.3. Observations on VBM analysis for ASD in females

VBM analysis performed in Section 3.2 show that the motor cortex area is generally affected among female ASD patients. These observations have been summarized in Table 3. The primary motor cortex, premotor cortex, supplementary motor cortex and the somatosensory cortex are the sub-regions within the motor cortex that have shown activations in female ASD patients in our studies which is found to be consistent with the literature (Lai et al., 2013). The motor cortex area is responsible for planning, control and execution of movement. The same area is also responsible for

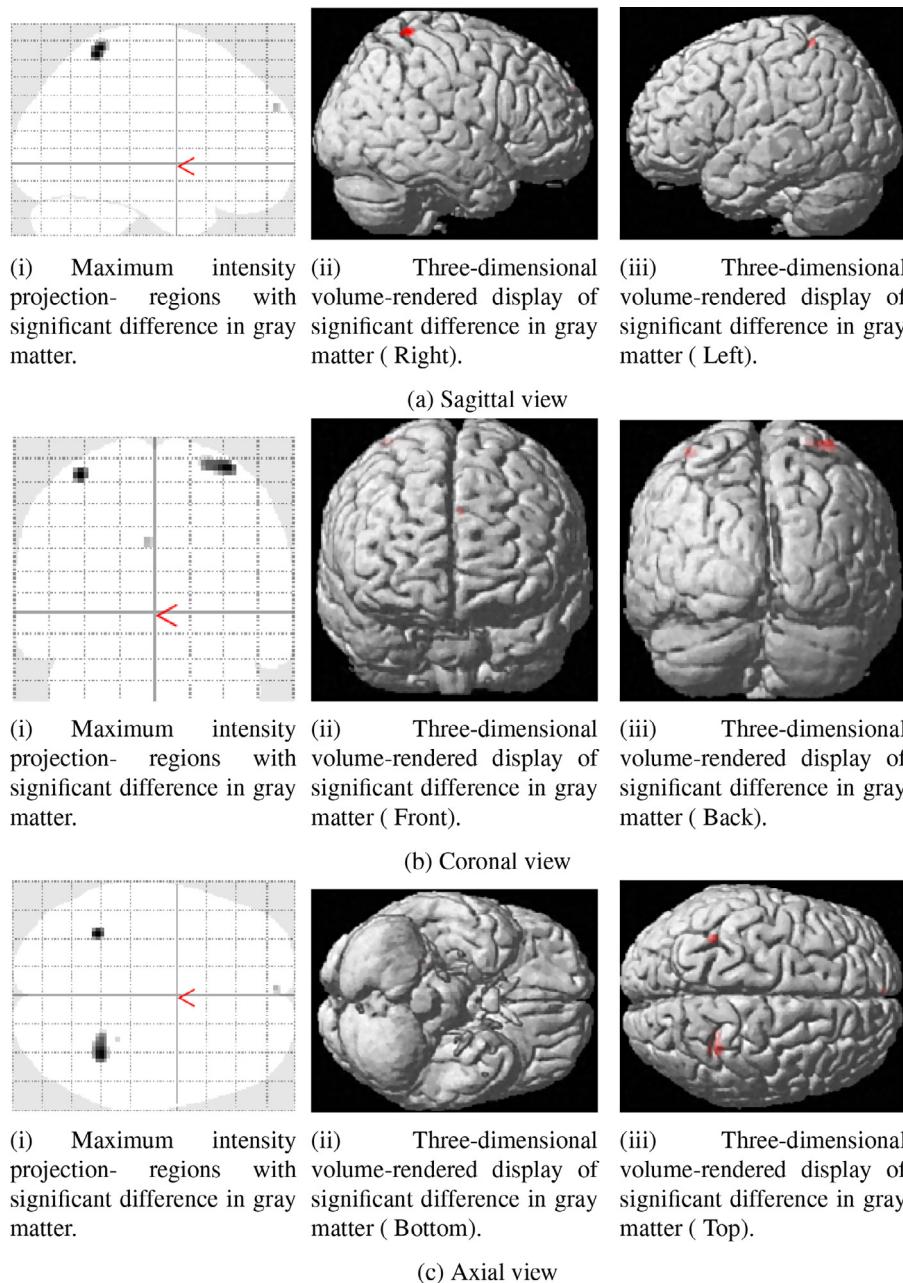


Fig. 5. Brain regions with significant differences in gray matter density in ASD patients, when considering MRI scans of adolescent females in the dataset.

understanding the actions of others by internally imitating the actions with the help of mirror neurons. The mirror neurons are active when the person performs the action and also when the person watches others performing the same action. One of the main characteristics of ASD patients is their difficulty in interpreting basic social cues such as pointing at objects, etc. Thus the identified regions seem to be consistent with this symptom. Apart from VBM, EEG and functional MRI studies have noted dysfunction of the mirror neuron system among ASD patients (Müller, Pierce, Ambrose, Allen, & Courchesne, 2001; Oberman et al., 2005).

However, different age groups show activations for different sub-regions within the motor cortex. For example, the somatosensory cortex is affected for adolescents while adults show activation in the premotor cortex and the supplementary motor cortex. When considering the whole set of females, the primary motor cortex is affected and additionally, the paracentral lobule region also shows some activation.

3.4. VBM analysis for ASD detection in males

In this section, the results of the VBM analysis for males i.e the three Cases D, E and F mentioned in Section 3.1 are provided and discussed. For each case, the maximum intensity projection, the location of significant voxels and the number of features obtained are given.

3.4.1. Case D: considering all males

In this case, all the MRI scans belonging to males in the ABIDE dataset are taken and a VBM analysis is done on this whole set. The MIP and 3D rendering are shown for this case in Fig. 7.

From Fig. 7, it can be observed the lateral ventricle, caudate, pre-motor cortex, supplementary motor cortex and the medial frontal gyrus regions of the brain, show significant difference in gray matter composition for ASD patients. From these voxel locations, the gray matter tissue probabilities are taken as the features resulting in 645 features for each MRI scan.

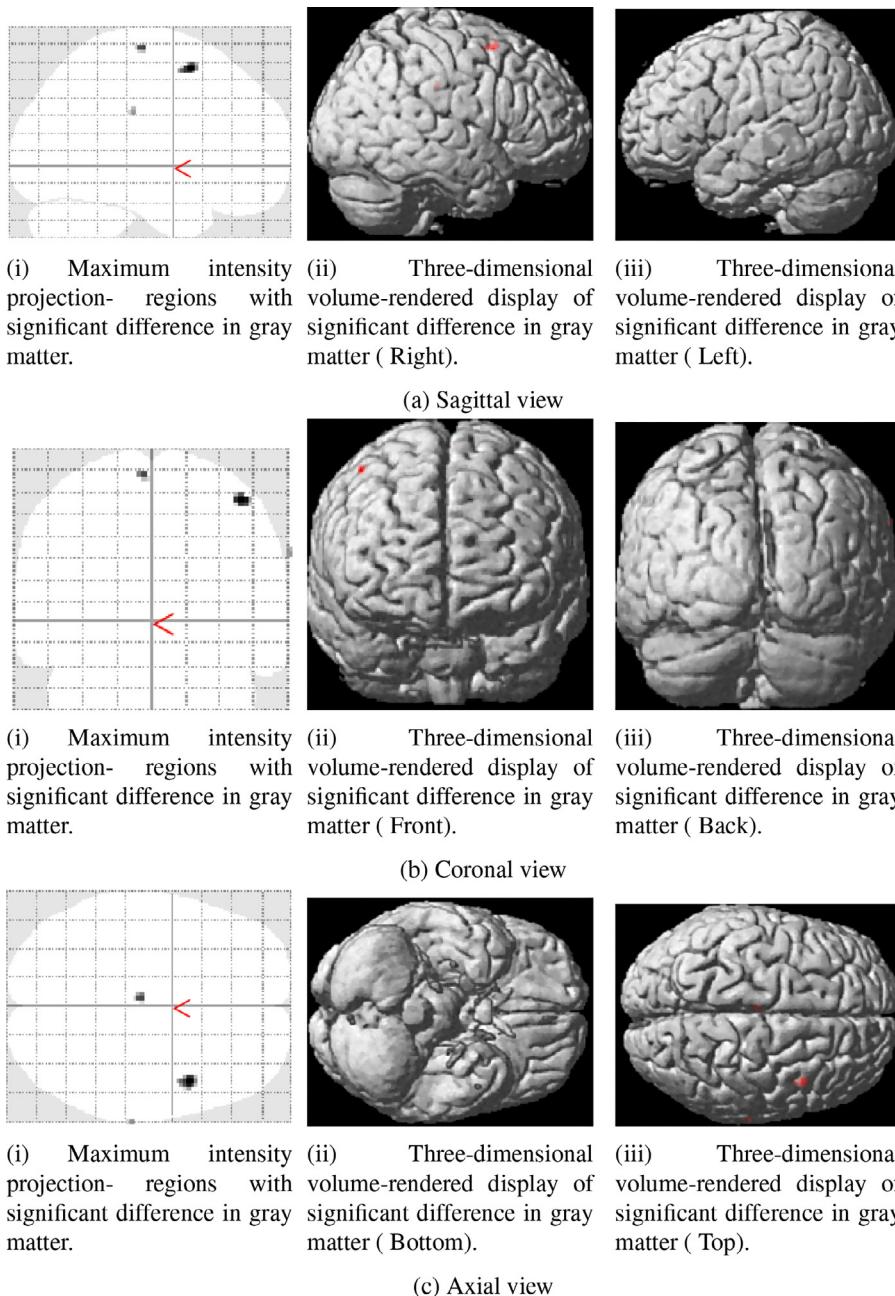


Fig. 6. Brain regions with significant differences in gray matter density in ASD patients, when considering MRI scans of adult females in the dataset.

3.4.2. Case E: ASD detection for adolescent males (age <18 years)

In this case, all the MRI scans belonging to adolescent males (<18 years) in the ABIDE dataset are considered separately for VBM analysis. The MIP and the 3D rendering of the activated areas are shown in Fig. 8 for this case. From this figure, it can be seen that the precentral gyrus, paracentral lobule, medial frontal gyrus, primary somatosensory cortex, primary motor cortex, premotor cortex and supplementary motor cortex regions are affected for adolescent males. The VBM analysis of MRI from adolescent males resulted in 2699 features for each MRI scan.

3.4.3. Case F: ASD detection for adult males (age > = 18 years)

In this case, the MRI scans from the male adults (Age > = 18 years) in the ABIDE dataset are considered separately. VBM analysis is performed as done in the previous sections in order to identify the brain regions that show significant differences in the gray matter compo-

sition for adult males suffering from ASD. The maximum intensity projections ($P<0.001$) and the 3-dimensional rendering of activated areas for this case is presented in Fig. 9. From this figure, it can be seen that the superior frontal gyrus, frontal eye fields areas are affected for adult males. 132 features are obtained for each MRI scan by considering the gray matter probability in these areas.

3.5. Observations on VBM analysis for ASD in males

The regions affected for males of different age-groups have been summarized in Table 3. When considering all the males in the dataset (Case D), the pre-motor cortex, supplementary motor cortex, caudate, lateral ventricle and the medial frontal gyrus regions are the areas identified by VBM. This has considerable overlap with the regions identified for adolescent males (Case E) where the motor cortex, medial frontal gyrus and the paracentral lobule regions are

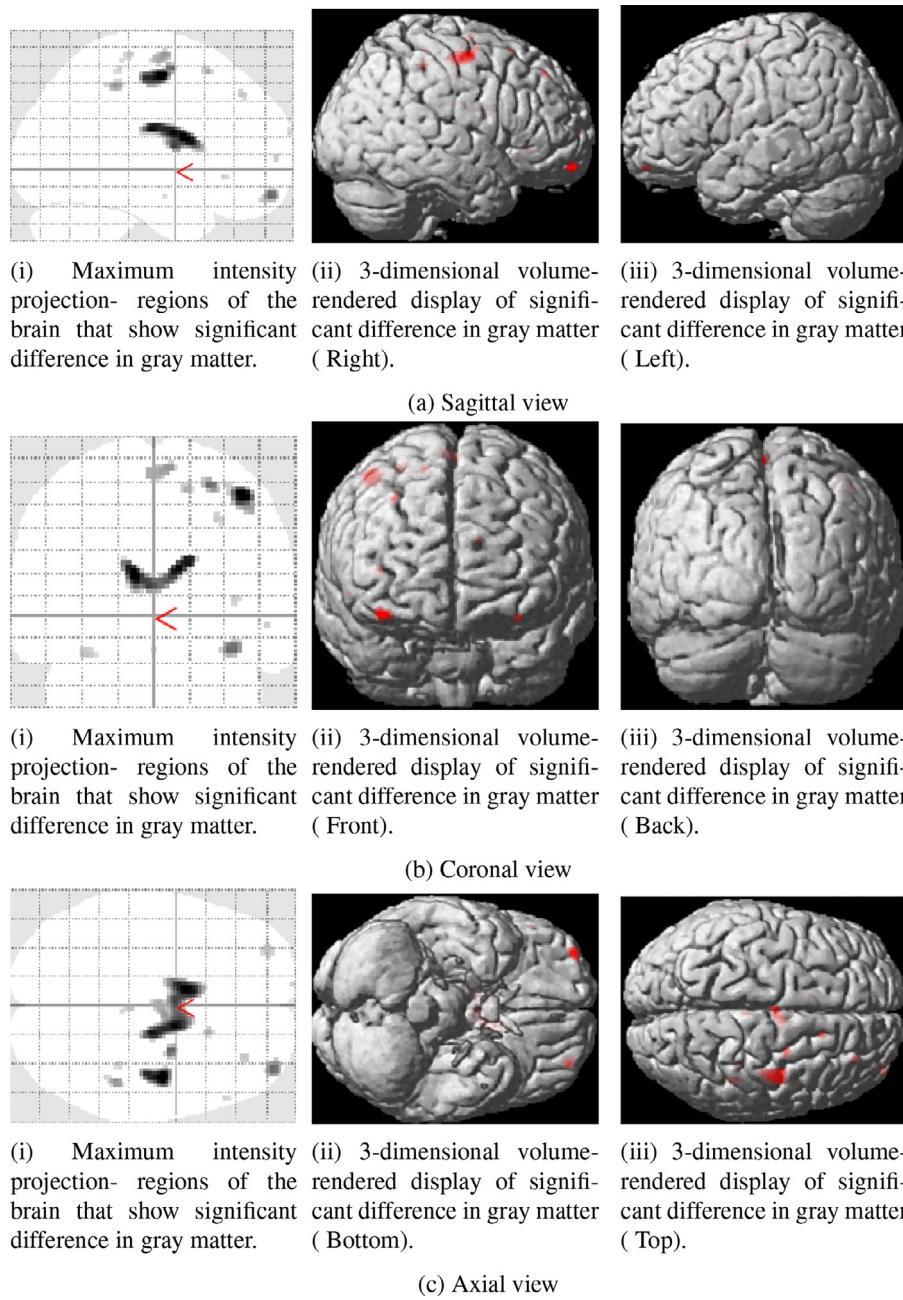


Fig. 7. Brain regions with significant differences in gray matter density in ASD patients, when considering MRI scans of all males in the ABIDE dataset.

identified by VBM. It can be noted that the motor cortex and the paracentral lobule are also affected for females as discussed in the earlier section where its connection to understanding actions of others was explained. However, for male adults (Case F), these regions do not seem to be significantly affected. Adults show activation in the superior frontal gyrus and the frontal eye fields area of the brain. These regions are responsible for self-awareness and control of visual attention etc. Thus wide variations are observed in the brain regions affected for adolescent and adult males.

It is known from recent longitudinal studies on autism that the anatomic pathology of autism changes with age and sometimes it leads to age-specific abnormalities (Courchesne et al., 2011a; Courchesne et al., 2011b). While continuous brain changes were found to occur from childhood to old age, the rates and types of change were found to be likely to vary with brain region. These lon-

gitudinal studies have also found that it is likely that there are age-specific defects in the brain implying that the defects present at adolescence may not be present at adulthood and vice versa. Nonetheless, potentially some subset of neural defects may also be present across a wide age range (Courchesne et al., 2011a). It can be observed that when considering all males and adolescent males, the features were picked up from a wider range of regions resulting in significantly higher number of features than the case for adult males. This is due to the larger variations in the adolescent male brain, which is undergoing rapid development following a unique developmental path. Even among females, VBM resulted in features from wider regions for adolescents than adults. Thus by isolating adults from adolescents and studying them separately, a clearer picture of the affected regions can be obtained. Thus the VBM results presented in this study also validates the theory of age-specific anatomic abnormalities in autism.

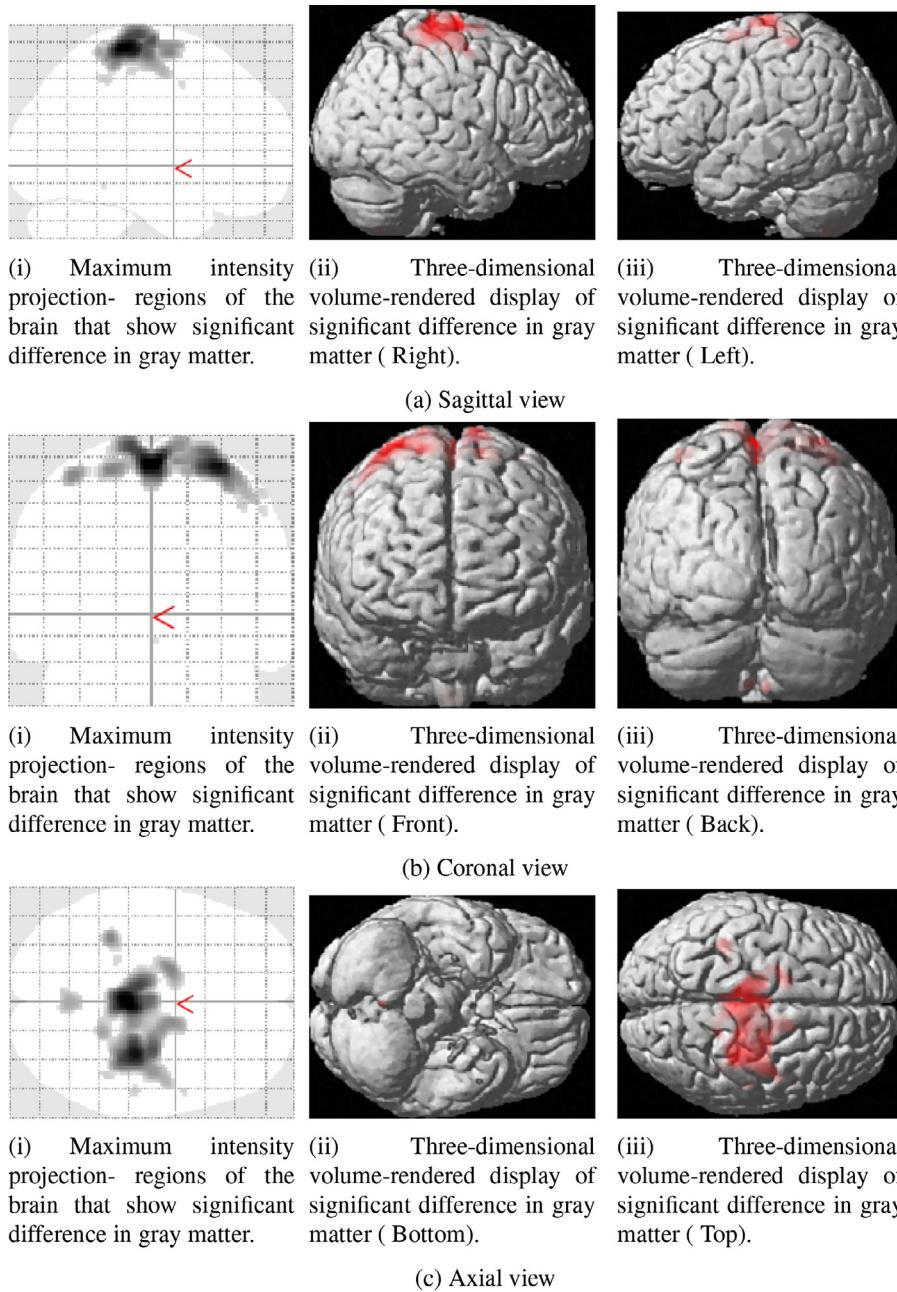


Fig. 8. Brain regions with significant differences in gray matter density in ASD patients, when considering MRI scans of adolescent males in the ABIDE dataset.

3.6. Classification performance comparison

Using the features obtained from VBM analysis, the performance of the EMcRBFN classifier is compared with the SVM, MLP, ELM and PBL-McRBFN classifiers for the ASD detection problem and presented in this section. The classification results are presented for the six cases that were considered in [Section 3.1](#), in the same order.

3.7. ASD classification performance for females

In this section, the ASD classification studies are done exclusively for females based on the features obtained from VBM for Cases A, B and C as described earlier. For each of these cases separate classifiers are built using the five algorithms mentioned above and their overall classification accuracy and F-measure are compared.

3.7.1. Case A: considering all females

Here, the classification results are presented for the case where all the MRI belonging to females are analyzed together. The features obtained in [Section 3.2.1](#) was used to build a classification model using the proposed EMcRBFN classifier and compared with the other classifiers like SVM, MLP, ELM and PBL-McRBFN is given in [Table 4](#). The mean and the standard deviation of the overall accuracy and the F-measure for the various classifiers are compared for both the training and the testing phase. For each entry, in the table, the mean value obtained over the 10 trials are given first, followed by the standard deviation which is provided within parentheses. From the table, it can be seen that all the three recent neural network based approaches (ELM, PBL-McRBFN, and EMcRBFN) achieve better testing accuracy and F-measure than the popular SVM, and the MLP network achieves the least performance among all the classifiers. Number of neurons used by ELM, PBL-McRBFN and EMcRBFN are just about half of the

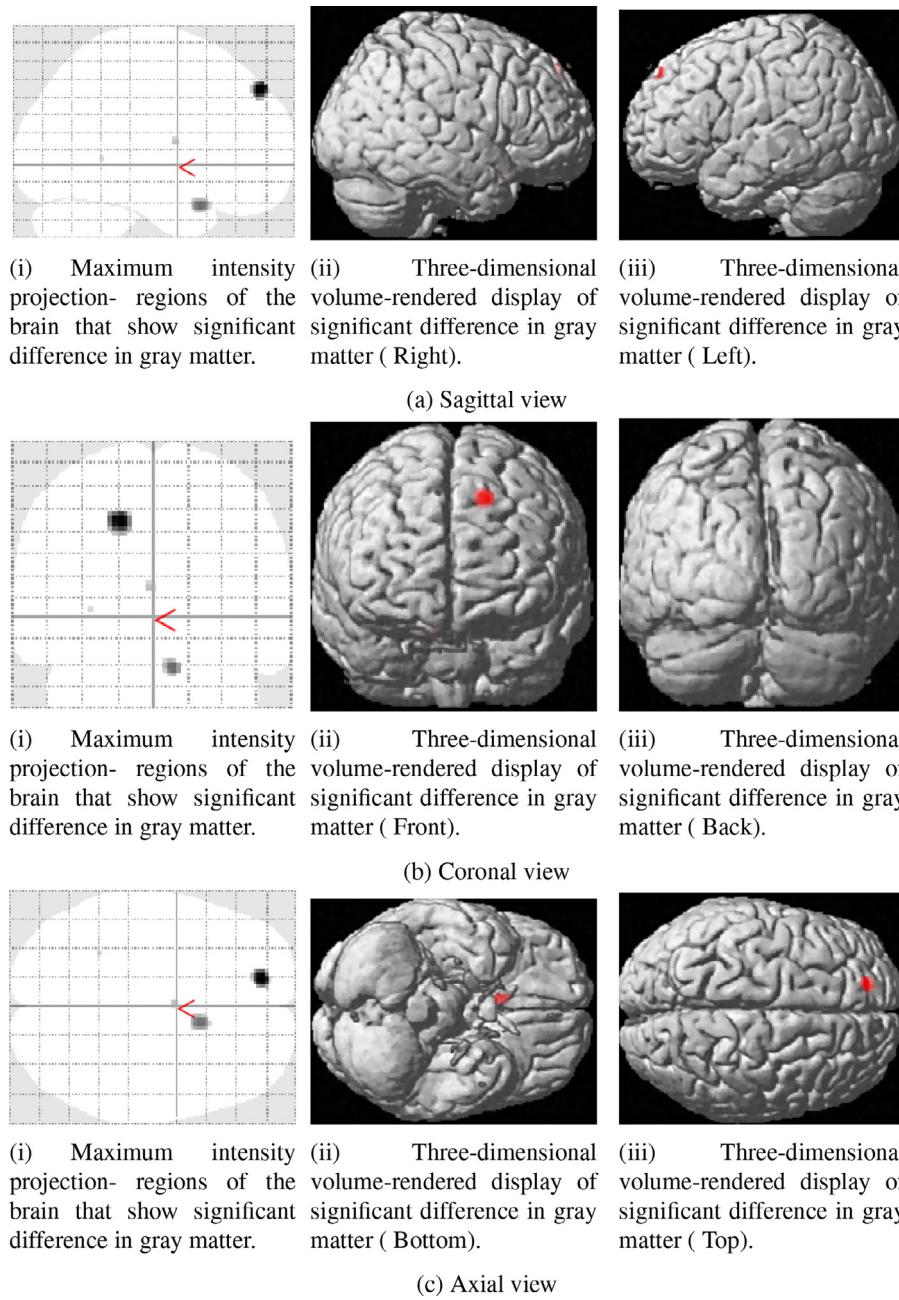


Fig. 9. Brain regions with significant differences in gray matter density in ASD patients, when considering MRI scans of adult males in the ABIDE dataset.

number of support vectors needed for SVM. Between ELM, PBL-McRBFN, and EMcRBFN, the lowest accuracy and F-measure is given by ELM while using more number of neurons. EMcRBFN gives the highest accuracy and F-measure among all the methods, however, it uses slightly more neurons than PBL-McRBFN. EMcRBFN is able to achieve approximately 3% improvement in accuracy and F-measure than PBL-McRBFN while using about 5 neurons more than PBL-McRBFN. At the same time, EMcRBFN is able to achieve about 10% improvement in F-measure and about 15% improvement in the overall accuracy than SVM. The training performance of all the algorithms except the MLP are approximately the same in terms of the accuracy and F-measure.

3.7.2. Case B: ASD detection for adolescent females (age < 18 years)

Here, the feature vectors obtained in Section 3.2.2 were used to build classification models for ASD detection. The classification per-

formance of EMcRBFN, SVM, MLP, ELM and PBL-McRBFN classifiers for female adolescents are given in Table 5. Overall, it can be seen that all the classifiers achieve better accuracy and F-measure for this case than the Case A, where MRI scans of all females were put together. Again, it can be noted that ELM, PBL-McRBFN and EMcRBFN classifiers provide much better overall accuracy and F-measure than the commonly used SVM method and MLP. Even during the training phase, SVM and MLP give lower accuracy in spite of using more number of support vectors/neurons. Further, the standard deviation of the accuracy and the F-measure are higher for the SVM than the neural network classifiers, both during the training and the testing phase. EMcRBFN achieves the highest testing accuracy and F-Measure. PBL-McRBFN is able to achieve slightly better testing accuracy and F-measure than ELM, but it uses 10 neurons lesser than ELM. EMcRBFN is able to see an improvement of about 2 to 3% in the testing accuracy and F-measure respectively, when compared to PBL-McRBFN, but it

Table 3

VBM detected regions in males for different age-groups.

Age-group	Identified regions	Functionality
Adolescent	Primary motor cortex, premotor cortex and supplementary motor cortex, primary somatosensory cortex, precentral gyrus, paracentral lobule medial frontal gyrus	Planning, controlling, and executing voluntary movements, execution of movement, executing complex movements, touch, motor and sensory stimulations, decision making
Adult	Superior frontal gyrus, frontal eye fields	Self-awareness, control of visual attention
All	Lateral ventricle, caudate, premotor cortex and supplementary motor cortex, medial frontal gyrus	Directed movements and mnemonic processing, planning, controlling, and executing voluntary movements, executing complex movements, decision making

Table 4

Classification performance of algorithms, considering all females regardless of age.

Algorithm	Training %		Testing %		
	Accuracy	F-measure	Accuracy	F-measure	Neurons
EMcRBFN	95.34(1.1)	94.2(1.3)	83.24(4.6)	75.89(7.7)	51.2(1.68)
PBL-McRBFN	95.86(1.6)	94.84(1.9)	80.27(5.6)	72.68(7.8)	46.1(6.3)
ELM	95.94(0.4)	94.67(0.6)	76.76(1.4)	68.42(3.4)	58(1.7)
SVM	95.01(9.8)	95.06(8.3)	72.43(5.4)	60.83(7.8)	93.7(11.2)
MLP	58.79(1.51)	49.92(4.95)	57.03(1.99)	50.44(4.64)	109(26.85)

Table 5

Classification performance of algorithms, considering only adolescent females.

Algorithm	Training %		Testing %		
	Accuracy	F-measure	Accuracy	F-measure	Neurons
EMcRBFN	96.92(2.9)	96.3(3.3)	85.86(2.5)	81.1(2.6)	40.7(4.2)
PBL-McRBFN	97.13(2.8)	96.24(3.4)	83.79(2.8)	78.36(4.6)	39.8(4.02)
ELM	96.67(1.0)	95.5(1.4)	83.44(2.2)	77.49(3.04)	50(3.3)
SVM	89.89(8.4)	84.2(13.6)	75.2(2.2)	62.3(5.98)	73.6(8.3)
MLP	61.03(1.75)	58.74(4.71)	59.66(2.84)	58.23(3.47)	81.4(22.6)

Table 6

Classification performance of algorithms, considering only adult females.

Algorithm	Training %		Testing %		
	Accuracy	F-measure	Accuracy	F-measure	Neurons
EMcRBFN	98.66(1.7)	98.77(1.99)	98.75(3.95)	98(6.3)	13.2(4.4)
PBL-McRBFN	98.67(1.7)	98.33(3.5)	96.25(6.03)	94(9.7)	13(2.45)
ELM	99(1.6)	98.77(1.99)	96.25(6.03)	94.57(8.8)	14.3(1.34)
SVM	95.33(6.5)	93.84(8.78)	87.5(5.9)	82.4(8.3)	26.4(4.5)
MLP	66.0(2.11)	61.86(5.44)	62.5(1.2)	59.61(10.1)	80(35.84)

uses slightly higher number of neurons. During the training phase, ELM, PBL-McRBFN and EMcRBFN achieve very similar accuracy and F-Measure. Among all the classifiers, MLP gives the least accuracy and F-measure both during testing and training.

3.7.3. Case C: ASD detection for adult females (age > = 18 years)

In this case, ASD classification performance is studied for female adults. The feature vectors obtained in Section 3.2.3 were used to build the five classifiers for ASD detection and their overall accuracy and F-measure are summarized in Table 6. It can be observed here that ELM, PBL-McRBFN and EMcRBFN are able to achieve high accuracy (>96%) and F-measure for this age-group. The conventional SVM and MLP are also able to achieve much better accuracy and F-measure that was achieved in Cases A and B, but it's still significantly lower (by about 10% and 30%, respectively) than the performance of

Table 7

Case D - classification performance of algorithms, considering all males regardless of age.

Algorithm	Training %		Testing %		
	Accuracy	F-measure	Accuracy	F-measure	Neurons
EMcRBFN	70.37(0.92)	70.11(2.77)	61.38(1.84)	61.84(4.83)	155(11.79)
PBL-McRBFN	74.36(4.89)	74.24(6.21)	59.64(2.41)	59.23(5.33)	168.1(39.26)
ELM	75.48(0.51)	75.28(0.49)	59.33(1.43)	58.58(1.84)	185(7.07)
SVM	89.45(0.71)	89.52(0.83)	55.19(2.6)	58.86(2.2)	451.2(107.7)
MLP	54.33(0.74)	53.81(2.39)	53.35(0.48)	53.29(3.10)	314(88.72)

ELM, PBL-McRBFN and EMcRBFN. SVM also uses nearly double the number of support vectors than the number of neurons used by ELM, PBL-McRBFN and EMcRBFN. EMcRBFN classifier is able to achieve 2% more accuracy and nearly 4% more in terms of F-measure than ELM. ELM is able to achieve slightly higher F-measure than PBL-McRBFN for this age group.

3.8. Observations on classification performance for females

Based on the classification results presented in Section 3.7, it can be seen that the classification accuracy for all the classifiers can be improved by performing separate VBM feature extraction and building separate classifiers for adults (Case C) and adolescents (Case B) instead of considering everybody as a single set (Case A). In particular, the classification accuracy for adults increased significantly when considering them separately. This may be due to the fact that it is easier to spot variations in brain structure among adults whose brain is supposed to be fully developed whereas the brain regions of adolescents may not be fully developed and take their own individual developmental path. Hence it is possible that variations can be observed even between normal adolescents and this makes it difficult to determine whether a particular variation is due to ASD. Among the classifiers, EMcRBFN outperforms the conventional SVM based approach both in terms of accuracy and in terms of F-measure and it is able to achieve an improvement of about 3% in classification accuracy over PBL-McRBFN. However, it has a tendency to consume a slightly higher number of neurons than PBL-McRBFN. The high classification accuracies achieved for adult and adolescent females show that the gray matter based features detected by VBM are sufficient for accurate classification of ASD among females.

3.9. ASD classification performance for males

In this section, the ASD classification studies are done exclusively for males based on the features obtained from VBM for Cases D, E and F. For each of these cases, separate classifiers are built using the five classifiers mentioned and their overall classification accuracy and F-measure are compared.

3.9.1. Case D: considering all males

In this section, the classification results are presented for the case where MRI belonging to all the males in the ABIDE dataset are analyzed together. The features obtained in Section 3.4.1 was used to build a classification model using the five classifiers. The overall accuracy, F-measure and the number of neurons used by each of the five classifiers is given in Table 7. From the table, it can be seen that the accuracy and F-measure of all the classifiers are significantly lower than that was achieved for females (Case A). There is just little difference in the classification accuracy between ELM and PBL-McRBFN, while EMcRBFN is able to achieve an improvement of 2.5% in classification accuracy over PBL-McRBFN and use about 15 neurons lesser than PBL-McRBFN. It should be noted that in this case 645 features were obtained which is significantly higher than the number of features obtained for females. Thus it can be seen that in such a

Table 8

Case E - classification performance of algorithms, considering only adolescent males.

Algorithm	Training %		Testing %		
	Accuracy	F-measure	Accuracy	F-measure	Neurons
EMcRBFN	84.28(7.25)	84.01(8.09)	61.81(5.7)	61.85(2.62)	192.3(57.83)
PBL-McRBFN	88.63(13.65)	88.48(14.11)	60.67(2.62)	61.28(5.19)	208.6(86.32)
ELM	85.83(0.69)	86.34(0.69)	58.51(3.31)	59.40(4.47)	218(4.21)
SVM	96.46(4.6)	96.73(4.22)	55.95(4.3)	59.97(5.8)	437(15.6)
MLP	54.7(0.89)	55.51(1.5)	53.58(0.64)	53.32(3.59)	290(40.28)

Table 9

Case F - classification performance of algorithms, considering only adult males.

Algorithm	Training %		Testing %		
	Accuracy	F-measure	Accuracy	F-measure	Neurons
EMcRBFN	83.83(4.04)	82.49(6.47)	70.55(1.74)	69.46(5.61)	103.3(6.78)
PBL-McRBFN	89.09(7.02)	87.8(8.91)	68.63(3.84)	68.08(3.0)	102.1(16.92)
ELM	85.82(1.17)	84.37(1.42)	67.81(1.86)	64.75(3.12)	105(5.27)
SVM	88.91(14.9)	87.96(16.44)	59.45(1.47)	48.96(8.55)	203(22.09)
MLP	56.55(1.12)	54.76(1.17)	55.62(0.96)	55.74(3.41)	132(44.04)

high-dimensional feature space, the q -Gaussian activation function of the EMcRBFN results in a performance improvement. The classification accuracy of SVM is about 6% lower than that of EMcRBFN and the accuracy of MLP is even lower.

3.9.2. Case E: ASD detection for adolescent males (age < 18 years)

Here, the feature vectors obtained in Section 3.4.2 were used to build classification models for ASD detection specific to adolescent males. It can be seen from Table 8 that the classification accuracies did not show any significant change than the previous case where all the males were considered together. Thus, it can be said that the strategy of building age-group specific classifier did not really improve the ASD detection accuracy for adolescent males. It should be noted here that this is the case with significantly high number of features (2699) than others and it is possible that the feature obtained from VBM are not discriminative enough to result in effective classification. However, as observed in the previous case, in such high-dimensional feature space, EMcRBFN is able to achieve a slightly better classification accuracy than PBL-McRBFN while using lesser neurons.

3.9.3. Case F: ASD detection for adult males (age > = 18 years)

In this case, ASD classification performance is studied for adult males by using the feature vectors obtained in Section 3.4.3 to build classification models. From Table 9, it can be observed that all the classifiers are able to achieve much higher accuracy for this age-group than what was observed from the previous two cases. The conventional SVM is also able to achieve 4% better accuracy than that was achieved in Cases D and E, but its accuracy is still about 10% lower than the performance of EMcRBFN. EMcRBFN is able to achieve 2.5% better accuracy than PBL-McRBFN but in this case it uses slightly higher number of neurons.

3.10. Observations on classification performance for males:

The classification accuracies achieved for males is also significantly lower by about 25% than the corresponding classifiers for females. However, even among males, it can be seen that the classification performance for all the classifiers can be improved by performing separate VBM feature extraction and building separate classifier for adults (Case F) instead of considering all the males as a single set (Case D). But for the case of male adolescents (Case E), age-group specific feature extraction and classification does not have any significant impact on the classification accuracy. Based on the classification results and the larger number of features it can be said that the

approach presented in this paper is not sufficient for accurate detection of ASD in males even though the proposed EMcRBFN classifier is able to achieve higher classification accuracy than the commonly used SVM. As in the case of females, lower classification accuracy was obtained for adolescents, than adults and features from wider regions of the brain were picked up by VBM for adolescents than adults. This again shows that there is high variation in the gray matter based features among adolescents with a developing brain making it difficult to find a functional relationship between such variations and ASD. Nevertheless, it is clear from these studies that by considering adults separately, this problem can be avoided for them resulting in better ASD detection performance.

4. Conclusions

This paper has presented an approach for accurate detection of ASD from structural MRI of the brain with an extended metacognitive radial basis function neural classifier. The extended McRBFN classifier is able to handle the high dimensional features generated by the whole brain VBM approach and approximates the classifier functional relationship very effectively. Using the publicly available ABIDE dataset, the performance of the proposed approach has been compared with the state-of-the art classifiers (SVM, ELM and MLP). The q -Gaussian activation function and the modified learning algorithm helps the EMcRBFN classifier to achieve 2–12% improvement in classification accuracy over SVM, which is commonly used for ASD detection.

By taking a cue from the findings in medical literature that ASD manifests differently in males and females and also among adults and adolescents, a comprehensive analysis of ASD detection based on different gender (male/female) and age groups (adolescent/adult) is provided in this paper which validates the medical findings. By considering males and females separately it is found that ASD can be detected more accurately among females (81%) than in males (60%). Within each gender, it is found that the ASD detection performance can be improved by a further 10–15% for adults and about 1–3% for adolescents by performing separate VBM analysis and classification for these age groups. For the first time in the literature, such a large-scale, comprehensive analysis has been provided separately for gender (male/female) and age groups (adolescent/adult) by using the openly available ABIDE dataset for structural MRI based ASD detection.

Among females, the results indicate gray matter atrophy within different sub regions of motor cortex for ASD patients consistent with the medical literature. For adolescent females, somatosensory cortex region is affected which could be responsible for inability to identify objects by touch while among adult females, premotor cortex and supplementary motor cortex are affected which may be responsible for an impaired or inability to plan, control, and execute voluntary movements. High classification accuracies of 98 and 83% have been achieved for adult and adolescent females indicating that the features identified by VBM based on gray matter atrophy are sufficient for accurate ASD detection among females.

Among males, differences in gray matter composition has been found in many other regions of the brain apart from the motor cortex area, thus resulting significantly higher dimensions for the feature vector. For adolescent males with ASD, the primary motor cortex, premotor cortex, supplementary motor cortex, primary somatosensory cortex, precentral gyrus, paracentral lobule and the medial frontal gyrus are found to be affected which may lead to an inability to plan, control, execute voluntary movements, complex movements, decision making and identifying objects by touch. Among male adults with ASD, superior frontal gyrus and frontal eye fields are affected which may be responsible for lack of self awareness and visual attention. However, the low detection accuracies (60% for adolescent

males, 70% for adult males) and the variation in the regions identified suggest that further studies are needed.

The low classification performance for adolescents could be due to anatomical variability between them since their brain is still undergoing fast development following a unique developmental path. Hence, the use of age and gender specific templates for VBM analysis needs to be explored. Further work is also needed for understanding the severity of ASD, for which one needs to study the subcategory (autism, asperger's etc) of ASD and the specific regions where gray matter composition is affected for these subcategories for better diagnosis.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions, which improved the quality of the paper.

References

- ABIDE (2013). Autism brain imaging data exchange, http://fcon_1000.projects.nitrc.org/indi/abide/.
- APA (2013). *The diagnostic and statistical manual of mental disorders: DSM 5*. US: Book point.
- Ashburner, J., & Friston, K. (2005). Unified segmentation. *NeuroImage*, 26, 839–851.
- Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry – the methods. *NeuroImage*, 11(6), 805–821.
- Babu, G. S., & Suresh, S. (2013). Sequential projection-based metacognitive learning in a radial basis function network for classification problems. *IEEE Transactions on Neural Networks and Learning Systems*, 24(2), 194–206.
- Babu, G. S., Suresh, S., & Mahanand, B. (2013). Meta-cognitive q-gaussian RBF network for binary classification: application to mild cognitive impairment (MCI). In *Proceedings of the 2013 international joint conference on neural networks (IJCNN)* (pp. 1–8). IEEE.
- Boddaert, N., Chabane, N., Gervais, H., Good, C., Bourgeois, M., Plumet, M., et al. (2004). Superior temporal sulcus anatomical abnormalities in childhood autism: a voxel-based morphometry MRI study. *NeuroImage*, 23(1), 364–369.
- Bonilha, L., Cendes, F., Rorden, C., Eckert, M., Dalgarondo, P., Li, L. M., et al. (2008). Gray and white matter imbalance—typical structural abnormality underlying classic autism? *Brain and Development*, 30(6), 396–401.
- Calderoni, S., Retico, A., Biagi, L., Tancredi, R., Muratori, F., & Tosetti, M. (2012). Female children with autism spectrum disorder: an insight from mass-univariate and pattern classification analyses. *NeuroImage*, 59(2), 1013–1022.
- CDC (2014). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *Morbidity and Mortality Weekly Report. Surveillance Summaries* (Washington, DC: 2002), 63, 1.
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2, 27:1–27:27 Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- Courchesne, E., Campbell, K., & Solso, S. (2011a). Brain growth across the life span in autism: age-specific changes in anatomical pathology. *Brain Research*, 1380, 138–145.
- Courchesne, E., Webb, S., & Schumann, C. (2011b). From toddlers to adults: the changing landscape of the brain in autism. *Autism Spectrum Disorders*, 611–631.
- Ecker, C., Marquand, A., Mourao-Miranda, J., Johnston, P., Daly, E. M., Michael, J. B., et al. (2010a). Describing the brain in autism in five dimensions – magnetic resonance imaging – assisted diagnosis of autism spectrum disorder using multiparameter classification approach. *Journal of Neuroscience*, 30.
- Ecker, C., Rocha-Rego, V., Johnston, P., Mourao-Miranda, J., Marquand, A., Daly, E. M., et al. (2010b). Investigating the predictive value of whole-brain structural mr scans in autism: a pattern classification approach. *NeuroImage*, 40, 44–56.
- Eisenmajer, R., Prior, M., Leekam, S., Wing, L., Gould, J., Welham, M., et al. (1996). Comparison of clinical symptoms in autism and Asperger's disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(11), 1523–1531.
- Fernández-Navarro, F., Hervás-Martínez, C., Gutiérrez, P. A., Peña-Barragán, J. M., & López-Granados, F. (2012). Parameter estimation of q-gaussian radial basis functions neural networks with a hybrid algorithm for binary classification. *Neurocomputing*, 75(1), 123–134.
- Geschwind, D. H., & Levitt, P. (2007). Autism spectrum disorders: developmental disconnection syndromes. *Current Opinion in Neurobiology*, 17(1), 103–111.
- Hadjikhani, N., Joseph, R. M., Snyder, J., & Tager-Flusberg, H. (2006). Anatomical differences in the mirror neuron system and social cognition network in autism. *Cerebral Cortex*, 16(9), 1276–1282.
- Hardan, A. Y., Girgis, R. R., Lacerda, A. L., Yorbik, O., Kilpatrick, M., Keshavan, M. S., & Minshew, N. J. (2006). Magnetic resonance imaging study of the orbitofrontal cortex in autism. *Journal of Child Neurology*, 21, 866–871.
- Hardan, A. Y., Libove, R. A., Keshavan, M. S., Melhem, N. M., & Minshew, N. J. (2009). A preliminary longitudinal magnetic resonance imaging study of brain volume and cortical thickness in autism. *Biological Psychiatry*, 66(4), 320–326.
- Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification. Technical report, Department of Computer Science, National Taiwan University.
- Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, 70(1), 489–501.
- Hyde, K. L., Samson, F., Evans, A. C., & Mottron, L. (2010). Neuroanatomical differences in brain areas implicated in perceptual and other core features of autism revealed by cortical thickness analysis and voxel-based morphometry. *Human Brain Mapping*, 31(4), 556–566.
- Iidaka, T. (2015). Resting state functional magnetic resonance imaging and neural network classified autism and control. *Cortex*, 63, 55–67.
- Jacquemont, S., Coe, B. P., Hersch, M., Duyzend, M. H., Krumm, N., Bergmann, S., et al. (2014). A higher mutational burden in females supports a “female protective model” in neurodevelopmental disorders. *The American Journal of Human Genetics*, 94(3), 415–425.
- Lai, M.-C., Lombardo, M. V., Suckling, J., Ruigrok, A. N., Chakrabarti, B., Ecker, C., et al. (2013). Biological sex affects the neurobiology of autism. *Brain*, 136(9), 2799–2815.
- Lord, C. (1989). Autism diagnostic observation schedule: a standardised observation of communicative and social behaviour. *Journal of Autism and Developmental Disorders*, 19, 185–212.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 659–685.
- McAlonan, G. M., Cheung, V., Cheung, C., Suckling, J., Lam, G. Y., Tai, K., et al. (2005). Mapping the brain in autism. A voxel-based MRI study of volumetric differences and intercorrelations in autism. *Brain*, 128(2), 268–276.
- Müller, R.-A., Pierce, K., Ambrose, J. B., Allen, G., & Courchesne, E. (2001). Atypical patterns of cerebral motor activation in autism: a functional magnetic resonance study. *Biological Psychiatry*, 49(8), 665–676.
- Nielsen, J. A., Zielinski, B. A., Fletcher, P. T., Alexander, A. L., Lange, N., Bigler, E. D., et al. (2013). Multisite functional connectivity MRI classification of autism: abide results. *Frontiers in Human Neuroscience*, 7.
- Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., & Pineda, J. A. (2005). Eeg evidence for mirror neuron dysfunction in autism spectrum disorders. *Cognitive Brain Research*, 24(2), 190–198.
- Orru, G., Petterson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using support vector machine to identify imaging biomarkers for neurological and psychiatric disease: a critical review. *Neuroscience and Biobehavioral Reviews*, 36, 1140–1152.
- Rojas, D. C., Peterson, E., Winterrowd, E., Reite, M. L., Rogers, S. J., & Tregellas, J. R. (2006). Regional gray matter volumetric changes in autism associated with social and repetitive behavior symptoms. *BMC Psychiatry*, 6(1), 56.
- Sateesh Babu, G., & Suresh, S. (2012). Meta-cognitive neural network for classification problems in a sequential learning framework. *Neurocomputing*, 81, 86–96.
- Sateesh Babu, G., & Suresh, S. (2013). Parkinson's disease prediction using gene expression – a projection based learning meta-cognitive neural classifier approach. *Expert Systems with Applications*, 40(5), 1519–1529.
- Savitha, R., Suresh, S., & Kim, H. (2014). A meta-cognitive learning algorithm for an extreme learning machine classifier. *Cognitive Computation*, 6(2), 253–263.
- SPM8 (2011). Wellcome trust center for neuroimaging. Institute of Neurology, UCL, London, UK, <http://www.fil.ion.ucl.ac.uk/spm/>
- Vigneshwaran, S., Mahanand, B. S., Suresh, S., & Savitha, R. (2013). Autism spectrum disorder detection using projection based learning meta-cognitive RBF network. In *Proceedings of the 2013 international joint conference on neural networks (IJCNN)* (pp. 1–8). IEEE.
- Vigneshwaran, S., Suresh, S., Mahanand, B. S., & Sundararajan, N. (2015a). ASD detection in males using MRI- an age-group based study. In *Proceedings of the 2015 international joint conference on neural networks (IJCNN)*. IEEE.
- Vigneshwaran, S., Suresh, S., & Sundararajan, N. (2015b). Improved classification performance of q-gaussian meta-cognitive RBF classifier. In *Proceedings of the 2015 international conference on cognitive computing and information processing (CCIP)* (pp. 1–6). IEEE.
- Volkmar, F. R., Cohen, D. J., Bregman, J. D., Hooks, M. Y., & Stevenson, J. M. (1989). An examination of social typologies in autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, 28(1), 82–86.
- Waiter, G. D., Williams, J. H., Murray, A. D., Gilchrist, A., Perrett, D. I., & Whiten, A. (2004). A voxel-based investigation of brain structure in male adolescents with autistic spectrum disorder. *NeuroImage*, 22(2), 619–625.