Alzheimer's Disease Detection by Utilizing Key Slice Selection in 3D MRI Images

Masoud Moradi and Hasan Demirel
Department of Electrical and Electronic Engineering
Eastern Mediterranean University
Famagusta, via Mersin 10, Turkey
masoud.moradi@cc.emu.edu.tr

Pouya Bolourchi
Department of Electrical and Electronic Engineering
Final International University
Kyrenia, via Mersin 10, Turkey
pouya.bolourchi@final.edu.tr

***Abstract* - This study proposes a new approach for improving the accuracy of high-dimensional pattern recognition problem of Alzheimer's Disease (AD). The proposed method uses information from three-dimensional Magnetic Resonance Imaging (MRI) brain data with 2D slices in three orthogonal directions. It includes the calculation of Fisher Criterion between the AD and Healthy Control (HC) groups in order to select key-slices in the coronal, sagittal and axial directions. The preprocessing phase involves region detection to segment Region of Interest (ROI) based on Displacement Field (DF) method. Then we utilized energy, contrast and homogeneity metrics along with feature vectors generated by PCA and Probability Distribution Function (PDF) methods in feature extraction phase for each key-slice selected in the earlier phase. Features coming from each key slice are combined through feature fusion for improved accuracy. Experimental results show fusion method that used with brain mask give us the higher or comparable results compared with other feature extraction techniques in the literature.**

Keywords - component; Alzheimer's disease; MRI; region of interest; data fusion; classification,

I. INTRODUCTION

Alzheimer's disease (AD) is an illness that affecting mostly older people especially those in their 60's [1]. It is a progressive dementia, which causes changes in behavior, loss of memory, thinking and language skills [2]. The AD syndrome gradually worsen over time, eventually meddles with a patient's daily normal life, and eventually kills the affected victim and so far, no cure has been discovered. Symptoms of AD has become popular for researchers from everywhere throughout the world due to its significance and impact on society [3]. All through the cycle of AD, side effects of AD may turn out to be more serious. From 2006, the overall global population that suffered from AD is estimated to be 26.6 million [4] [5]. It is estimated by 2050 the population of people who will be affected and stricken by AD to increase to 106 million and around 40 percent of this cases require intensive care [6]. Therefore, there is a need to improve a new and reliable procedure for diagnosing AD, which is likewise fundamental for the cure and administration of expediting the disintegration of Alzheimer disease [7].

High dimensional classification is imperative technique to achieve high accuracy of numerous applications, particularly in automated detection of patient suffering from Alzheimer Disease. A number of neuroimaging studies have been introduced by several high-dimensional pattern recognition procedures [8-10].

A feature-extraction procedure from a greater-dimensional data classification is improvised to decrease the computational cost and improve the accuracy. This procedure is very successful in processing the conclusive results. In this paper, three most imperative and effective feature extraction methods have been presented to solve the problem of high-dimensional pattern recognition in Alzheimer Disease.

The main potential attainment of machine learning for automated detection of patients who endure AD is high dimensional classification technique. Feature selection and dimension reduction technique from high dimensional data is one of the most important topics in pattern recognition and data mining. Over the last few years, many studies successfully verified high-dimensional pattern recognition in a number of neuroimaging studies [11]. In this paper, we introduce a new way for feature extraction in high-dimensional detection of AD.

A major target of machine learning derived from high-dimensional detection procedures is a major automated classification of patients suffering from Alzheimer's disease (AD). The feature-extraction method from high-dimensional data is the biggest issue of the automated classification. In a number of neuroimaging studies, a number of researching studies investigated high-dimensional pattern detection approaches.

In this paper, we present imperative feature extraction procedures in high-dimensional recognition of Alzheimer Disease. Main contributions of this study can be listed as: Using fisher criteria for selecting key slices (KS) Generating 2D mask for segmentation of region of interest (ROI) and Introducing, a new method for feature extraction based on data fusion to develop classification enactment by merging PDF features, Eigen brain values and statistical features.

The reminder of this paper is listed as follows. In section II, we explain the proposed method, which it includes preprocessing, feature extraction, feature section and feature fusion. Section III contains the experimental results and discussions and finally we end up with conclusions.

0-7695-6405-4/18/$31.00 ©2018 IEEE
DOI 10.1109/UKSim.2018.00029

II. PROPOSED METHOD

For this section, a clear description of detecting AD from HC subjects will be fully outlined. In the first section of this paper, we gave a descriptive outline regarding the AD in detail and supported our discussion with preprocessing data downloaded from OASIS database containing cross-section of MRI data, longitudinal MRI data and additional data for each subject. Since the calculation region of interest for the whole brain is result depriving, a KS selection that will serve as the wholly scanner of the brain was introduced. The preprocessing section include region detection to section Region of Interest (ROI) based on Displacement Field (DF) method. In the next section, we call feature extraction we extract energy, contrast and homogeneity metrics along with feature vectors generated by PCA and Probability Distribution Function (PDF) methods for each key-slice. In the last section, features coming from each key-slice are combined through feature fusion for improved accuracy. In all implementation of propose method was drawn Fig. 1.

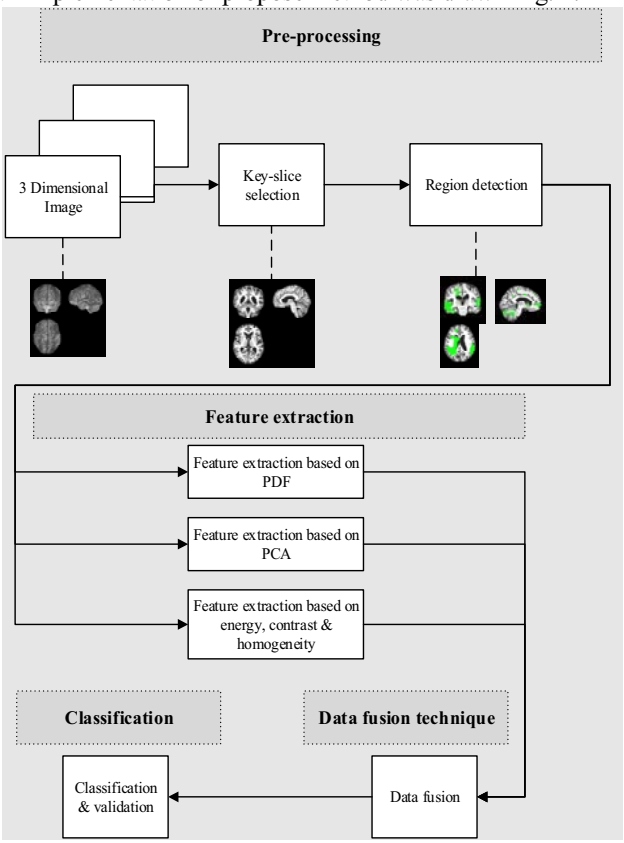

Figure 1. Block diagram of proposed AD detection method.

A. Key-Slice Selection

A key slice selection process which choose key slices having structures indicating AD from HC was introduced since the computation of displacement field on the wholly scanner of the brain was time-consuming [12]. A method based on fisher criterion J(w) to select the best key slices were presented as follows, given in (1):

$$J(w) = \frac{w^T S_B w}{w^T S_w w} \tag{1}$$

In (2) S_B is indicative of scatter matrix for classes L1 and L2 also in (3) S_w is indicative within class scatter matrix for classes L1 and L2 defined below as

$$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T \tag{2}$$

$$S_w = \sum_{x_i \in c_i} (x_i - \mu_{L1})(x_i - \mu_{L2})^T + \sum_{x_i \in c_i} (x_i - \mu_{L2})(x_i - \mu_{L2})^T \tag{3}$$

B. Region Detection

Firstly, we randomly selected one AD and one HC from our original data and we obtain DF for each key slice. The same process is repeated for all AD and HC samples in the train dataset (27 AD and 87 HC). This process is repeated because we have different regions for every sample from the dataset. We use all the images to obtain an average for every image. For the next step, we define threshold to generate the deformation map (region) in the brain

$$ROI = \left\{(x, y), \left|D(x, y)\right| \rangle T\right\} \tag{4}$$

where $D(x, y)$ represents the displacement field at the point of (x, y), $\left|\ \right|$ represents the magnitude and T represent the threshold. In other word, we consider the point with magnitude larger than T. According to the experiment, we consider value of T as five [13]. Smaller value of T may indicate more noises in the computed displacement field, whereas with large number of threshold we will lose realistic deformation. The process of ROI for each axis is shown by flowchart in Fig 5. In the first step, we calculate number of iterations. The number of iteration is equal to number of subject of ADs times number of subjects of HC. We then randomly selected one AD and one HC then we calculated DF for all the subjects and then we calculated the mean subjects in the last step according to the threshold we defined as ROI.

C. Feature Extraction

1) Feature Extraction Based on PDF: A PDF can be described as a vector of probabilities representing the probability of the pixel values and also defined as a raw feature vector extracted from image that belong into various disjointed intervals. It could also be defined as a statistical description of the distribution of occurrence probabilities of pixel values that can be considered a feature vector, defined as bins [11] [14]. The PDF extracted from voxels or pixels can be calculated as follows:

$$H = [p_1, p_2, ..., p_m], p_i = \frac{\eta_i}{N}, i = 1,2,...,m \tag{5}$$

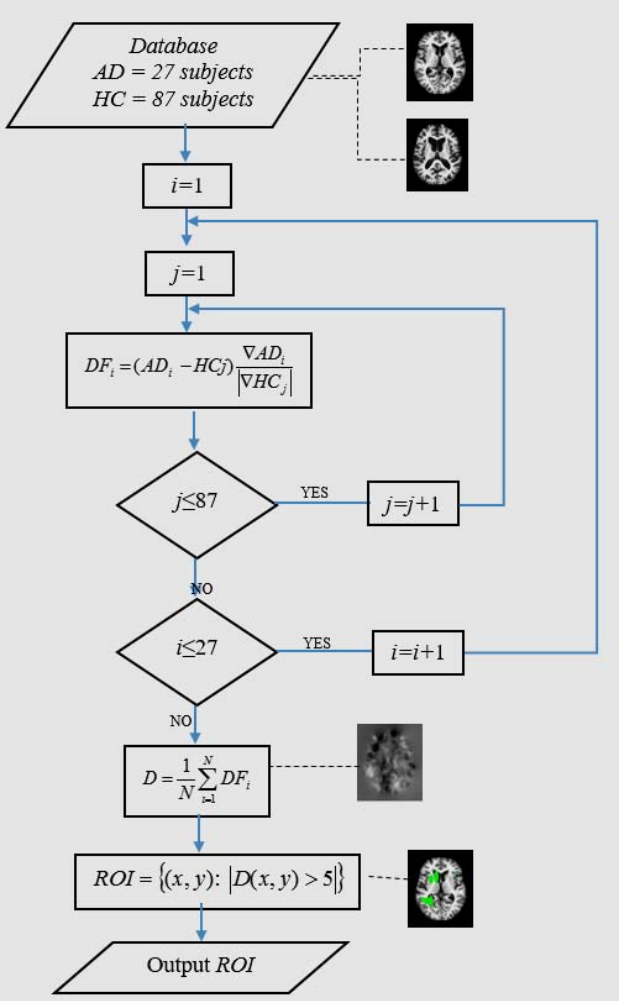

Figure 5. Flowchart of the proposed Region Detection Method

2) Feature Extraction Based on Statistical Parameters: Consider V represent two dimensional MRI data for each key-slice and V(x,y) as pixel located at (x,y). we utilized Energy E, Contrast C, Homogeneity H as the representation data V. They're expressed as*:*

$$E = \sum_{i,j} V(x,y)^2 \tag{6}$$

$$C = \sum_{i,j} |i-j|^2 V(i,j) \tag{7}$$

$$H = \sum_{i,j} \frac{V(i,j)}{1+|i-j|} \tag{8}$$

The triplet (energy, contrast, homogeneity) extracted from all key-slice of brain and then queued into a row vector.

3) Feature Extraction Based on PCA: Eigen-brain is a mathematical procedure that utilizes orthogonal advancement facilitated by PCA to match possibly corresponding and correlated variables into a set of uncorrelated values that could be defined as principal components [15]. In a case size M × N is derived from normalized dataset X, assuming M represents number of samples and N represents number of features. Normalized Z represents centered and scaled unit variance is derived by subtracting the mean value and dividing the resulting difference by its standard deviation value and thus this is represented by the formula below

$$Z = \frac{X - \mu(x)}{\sigma(x)} \tag{9}$$

The covariance of matrix C with size of N × N is expressed in the formula below

$$C = \frac{1}{M-1} ZZ^T \tag{10}$$

$M - 1$ instead of M is used to represent variance approximation value, Third, we express the Eigen decomposition of C:

$$C = U \Lambda U^{-1} \tag{11}$$

U is represented by N × (M − 1) matrix, whose confinements are the eigenvectors of matrix C covariance, and matrix 3 is represented by (M − 1) × (M −1) diagonal matrix whose diagonal values are eigenvalues of C, each correlating to an eigenvector N. It is usual process to sort the eigenvalue matrix 3 and matrix U by reducing eigenvalue $u_1 > u_2 > ... > u_N$. To view the i^{th} Eigen-brain u(i), the i^{th} column of U was reshaped to an image.

D. Data fusion

This part introduces data fusion technique to improve the accuracy of the proposed AD classification method. The goal of the data fusion technique is to combine the data from two or more distinct multiple source (vectors, classifiers) to improve performance [16].

E. Classification

The final step for distinguish AD from HC subjects is called Classification which evaluate performance serve as the last stage for distinguishing AD. In this paper, few classifiers have been introduced for AD classifications including Support Vector Machine (SVM) classifier, kNN and Decision Tree (DT) classifier.

III. RESULTS AND DISCUSSIONS

In this section, we discuss about results obtained from key-slice selection and feature extraction and classification. Finally, we compare our results with some state of the art methods.

A. Key-Slice Selection

The feature extraction approach on the whole brain was time consuming, and therefore we decided to select the best layer of brain in axial direction, sagittal direction and coronal direction based on fisher criteria rule. The optimal number

of slices are determined by fisher criteria. The fisher criteria between AD and HC group is calculated for all key slices and then selected as number of top discriminative key slices. In this study Fisher Criteria aid us to find the best key slices with the most discriminative pixels for the feature extraction and classification process. To find Fisher Criteria coefficient for each axis, we calculated the coefficient of fisher criteria for each slice. If the value of Fisher Criteria is equal to zero for each slice, the pixels of that layer does not have any difference between AD and HC, and the fact that the value of Fisher Criteria is high shows that these layers AD and HC differ from each other. The diagram for Fisher Criteria for Three axis are shown in Fig 2. For selection key slices in this axis, we choose an area that is higher than half the maximum value. As it is displayed in the diagram below, the area between 60 and 155 has more discriminative pixels compared to other slices. Since we want to have uncorrelated slices, we have to pick ten key slices from 60 to 150 with increasing step of 10. Figure 6.4 below shows key slices for coronal axis.

Figure 2. Curve of Fisher Criteria for coronal direction

The diagram of fisher criteria for sagittal axis is displayed in Fig 3, for selection key slices in this axis, we chose an area that is higher than half of the maximum value. As it is displayed in the diagram, the area between 40 and 130 has more discriminative pixels compare with other slices. Since we want to have uncorrelated slices, we have to pick ten key slices from 40 to 130 with increasing step of 10. Fig. 3 shows key slices for sagittal axis.

The diagram of fisher criteria for axial axis is shown in figure 4, for selection key slices in this axis, we chose an area that is higher than half the maximum value. In the diagram below, the area between 60 and 100 has more discriminative pixels compared to other slices. Since we want to have uncorrelated slices, we pick ten key slices from 60 to 100 with increasing step of 10. Fig 4 shows key slices for coronal axis.

Figure 3. Curve of Fisher Criteria for sagittal direction.

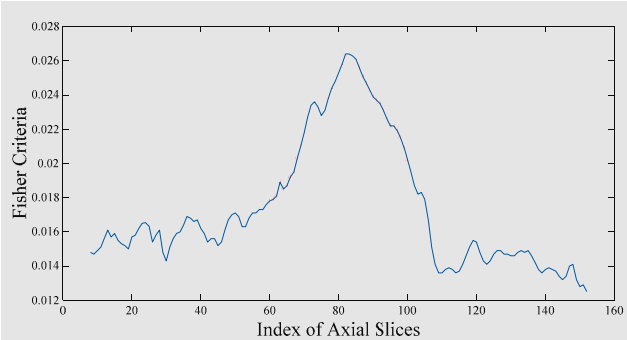

Figure 4. Curve of fisher criteria for axial direction.

B. Feature Extraction and Classification Comparison

In this section, we report accuracy of the 10-fold cross validation for detection of AD calculated over the OASIS data. In each axis for each slice we perform feature extraction based on PCA (10 features), PDF (20 features), Energy, Contrast and Homogeneity (ECH - 3 features). For each slice, 33 features were extracted. Coronal axis, sagittal axis and axial axis have 10 slices, 10 slices and 5 slices respectively. This means that 330 features are extracted from coronal and sagittal axis and 165 features are extracted from axial axis.

The experimental results show that the proposed methods are comparable with classical methods introduced in the literature. Table I shows the accuracy (%) of proposed method without using any masks. In this regard, we investigated the accuracy of different feature extraction techniques for each individual axis as Coronal, Sagittal, and Axial axis. The accuracy of extracted feature evaluated by utilizing different classifiers. It is clear that the PCA has the highest performance among all axis for kNN and SVM classifiers. The SVM classifier has the best performances when it applied to all feature extraction method in all axis. Furthermore, we used all three axes to evaluate the performance. As it is shown in Table I the accuracy is improved by having all three axes. Further improvement can be achieved if we fuse all axes. The highest accuracy is belonging to PCA using linear SVM with accuracy of 83.1%. The same approach can be done by utilizing mask as it is illustrated in Table II. The highest accuracy is achieved when

we applied fusion techniques followed by SVM classifier and it reaches to 88.1%.

TABLE I: THE ACCURACY OF THE PROPOSED METHOD WITHOUT USING MASK

Axis	**Feature extraction methods**	**classifier**		
		Decision Tree (%)	**Linear SVM (%)**	**kNN (%)**
Coronal axis	PCA	67.7	82.7	80.3
	PDF	68.5	75.6	72.4
	Energy, Contrast Homogeneity	73.2	78.0	78.0
Sagittal axis	PCA	73.2	82.1	81.1
	PDF	75.6	78.0	71.7
	Energy, Contrast Homogeneity	75.6	81.1	71.7
Axial axis	PCA	66.9	82.5	81.0
	PDF axial	73.2	79.5	77.2
	Energy, Contrast Homogeneity	77.2	79.5	70.1
Three axis	$PCA_{coronal}$ + $PCA_{sagittal}$+PCA_{axial}	71.7	**83.1**	**82.7**
	$PDF_{coronal}$ + $PDF_{sagittal}$+PDF_{axial}	69.3	76.4	74.8
	$ECH_{coronal}$ + $ECH_{sagittal}$+ECH_{axial}	70.9	77.2	74.8
PCA_{all} + PDF_{all} + ECH_{all}		**78.7**	82.7	80.3

TABLE II: THE ACCURACY OF THE PROPOSED METHOD WITH MASK

Axis	**Feature extraction methods**	**classifier**		
		Decision Tree (%)	**Linear SVM (%)**	**kNN (%)**
Coronal axis	PCA	76.4	85	80.3
	PDF	63.0	78.7	79.5
	Energy, Contrast Homogeneity	72.4	80.3	79.5
Sagittal axis	PCA	63	81.1	81.1
	PDF	70.1	84.3	77.2
	Energy, Contrast Homogeneity	69.3	79.5	76.4
Axial axis	PCA	64.6	73.2	75.6
	PDF	70.1	79.5	74.8
	Energy, Contrast Homogeneity	71.7	76.4	78
Three axis	$PCA_{coronal}$ + $PCA_{sagittal}$+PCA_{axial}	**80.3**	80.3	83.5
	$PDF_{coronal}$ + $PDF_{sagittal}$+PDF_{axial}	72.4	85	78
	$ECH_{coronal}$ + $ECH_{sagittal}$+ECH_{axial}	68.5	82.7	79.5
Proposed method (PCA_{all} + PDF_{all} + ECH_{all})		78.7	**88.1**	**85.1**

In Table III, we compared proposed method (PCA_{all} + PDF_{all} + ECH_{all} + LSVM) with other studies. Accuracy of the proposed method is higher than other methods, as well as, the sensitivity and specificity is comparable with the state of the art. The database used in this work, contains brain MRI scans of 30 AD and 97 HC subjects. The database adopted in this paper can be considered unbalanced by means of equal number of AD and HC samples. We had to use an unbalanced database due to difficulties faced to access private balanced AD databased. Despite of this disadvantage, the accuracy, sensitivity and specificity performances of the proposed methods are either higher or comparable with other methods.

TABLE III: COMPARISON OUR METHOD WITH OTHER METHODS

Method (Author)	Subjects (AD/HC)	ACC (%)	SEN (%)	SPE (%)
MSD + RBF SVM (Papakostas et al, 2015) [17]	49/49	85.00	78.00	92.00
VV + RBF-AB-SVM (Savio et al., 2011) [18]	49/49	86.00	80.00	92.00
GM+WM+SVM(Khedher et al., 2015) [19]	188/229	88.49	90.39	86.17
EB+WTT+RBF-KSVM (Dong et al., 2015) [15]	28/98	86.71	85.71	86.99
BRC + IG + VFI (Plant et al., 2010) [20]	32/18	78.00	65.63	100
Proposed method (PCA_{all} + PDF_{all} + ECH_{all} + LSVM)	30/97	88.10	80.76	91.08

IV. CONCLUSIONS

In this study, our focus was to propose a new approach to improve the accuracy of detection of Alzheimer's disease. In this regards, firstly we used key-slice in 3 dimension to reduce the dimensionality of MRI images. New feature extraction approach is proposed by utilizing 2D slices in three orthogonal directions. The proposed method includes the calculation of Fisher Criterion between the AD and HC groups in order to select key-slices in the coronal, sagittal and axial directions. Experimental results show the effectiveness of the proposed feature extraction with respect to the state-of-the-art. Extracted features coming from each key-slice are combined through feature fusion for improved accuracy. Experimental results show fusion method that used with brain mask generates higher or comparable results compared with other feature extraction techniques in the literature. Furthermore, we adopted three individual classifiers to compare our methods and the results show that SVM outperforms DT and K-NN classifiers. Finally, the proposed method achieved higher or comparable sensitivity, specificity and accuracy scores for the detection of AD among the other methods available in the literature.

REFERENCES

[1] R. J. Castellani, G. Perry, G. L. Iverson, "Chronic Effects of Mild Neurotrauma: Putting the Cart Before the Horse?," *Journal of Neuropathology & Experimental Neurology,* vol. 74, no. 6, p. 493–499, 2015.

[2] Y. Zhang, J. Yang, J. Yang, A. Liu and P. Sun, "A Novel Compressed Sensing Method for Magnetic Resonance Imaging: Exponential Wavelet Iterative Shrinkage-Thresholding Algorithm with Random Shift.," *International Journal of Biomedical Imaging,* vol. 2016, p. 10, 23 Feb 2016.

[3] R.Castellani, G.Perry , "The complexities of the pathology-pathogenesis relationship in Alzheimer disease.," *Elsevier,* vol. 88, no. 4, pp. 671-676, 2014.

[4] K. Hahn, N. Myers, S. Prigarin, K. Rodenacker, A. Kurz, H. Förstl, C. Zimmer, A.M. Wohlschläger and C. Sorg, "Selectively and progressively disrupted structural connectivity of func659tional brain networks in Alzheimer's disease Revealed by a novel framework to analyze edge distributions of networks detecting disruptions with strong statistical evidence.," *NeuroImage,* vol. 81, pp. 96-109, 2013.

[5] S. Goh, Z. Dong and Y. Zhang, "Mitochondrial Dysfunction as a Neurobiological Subtype of Autism Spectrum Disorder.," *JAMA Psychiatry,* vol. 71, pp. 665-671, 2014.

[6] M. Murphy, J. Huston, C. Jack, K. Glaser, A. Manduca, J. Felmlee and R.Ehman, "Decreased brain stiffness in Alzheimer's disease determined by magnetic resonance elastography.," *J Magn Reson Imaging,* vol. 34, no. 3, pp. 494-498, 2011.

[7] R. Brookmeyer, E. Johnson, K. Ziegler-Graham and H. Arrighi, "Forecasting the global burden of Alzheimer's disease. Alzheimers Dement.," *Alzheimer's & Dementia,* vol. 3, no. 3, pp. 186-191, 2007.

[8] M. Kantanen, S. Kiuru-Enari, O. Salonen, M. Kaipainen and L. Hokkanen, "Kantanen Subtle neuropsychiatric and neurocognitive changes in hereditary gelsolin amyloidosis (AGel amyloidosis).," *PeerJ,* vol. 2, p. e493, 2014.

[9] Y. Song and J. Wang, "Overview of Chinese research on senile dementia in mainland China.," *Ageing Research Review,* vol. 9, pp. S6-S12, 2010.

[10] Y. Zhang, Z. Dong, P. Phillips, S. Wang, G. Ji and J.Yang, "Exponential wavelet iterative shrinkage thresholding algorithm for compressed sensing magnetic resonance imaging.," *Information Sciences,* vol. 322, p. 115–132, 2015.

[11] I. Beheshti and H. Demirel, "Probability distribution function-based classification of structural MRI for the detection of Alzheimer's disease.," *Comput. Biol. Med,* vol. 64, p. 208–216, 2015.

[12] Q. Gao, J. Liu, H. Zhang, J. Hou, and X. Yang, "Enhanced fisher discriminant criterion for image recognition.," *Pattern Recognit,* vol. 45, no. 10, p. 3717–3724, 2012.

[13] Y. Zhang and S. Wang, "Detection of Alzheimer's disease by displacement field and machine learning.," *PeerJ,* vol. e1251, no. 3, 2015.

[14] H. Demirel and G. Anbarjafar, "Data fusion boosted face recognition based on probability distribution functions in different colour channels.," *EURASIP J. Adv Signal Process,* vol. 2009, 2009.

[15] Y. Zhang, Zh. Dong. Phillips, Sh. Wang, G. Ji and T. Yuan, "Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning.," *Frontiers in Computational Neuroscience,* vol. 9, p. 66:1–15, 2015.

[16] P. Bolourchi,M. Moradi,H. Demirel and S. Uysal , "Feature Fusion for Classification Enhancement of Ground Vehicle SAR Images.," in *UKSim-AMSS 19th International Conference on Modelling & Simulation*, 2017.

[17] G. a. Papakostas, a. Savio, M. Graña, and V. G. Kaburlasos, "A lattice computing approach to Alzheimer's disease computer assisted diagnosis based on MRI data.," *Neurocomputing,* vol. 150, pp. 37-42, 2015.

[18] A. Savio, M. T. García-Sebastián, D. Chyzyk, C. Hernandez, M. Graña, A. Sistiaga, A. López de Munain, and J. Villanúa, ""Neurocognitive disorder detection based on feature vectors extracted from VBM analysis of structural MRI.," *Comput. Biol. Med,* vol. 41, no. 8, pp. 600-610, 2011.

[19] L. Khedher, J. Ramírez, J. M. Górriz, a. Brahim, and F. Segovia, "Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images.," *Neurocomputing,* vol. 151, pp. 139-150, 2015.

[20] C. Plant,S.J Teipel,A. Oswald,C. Böhm,T. Meindl,J. Mourao-Miranda and M. Ewers, "Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease.," *Neuroimage,* vol. 50, no. 1, pp. 162-174, 2010.