Proc. of the Ninth Intl. Conf. on Advances in Bio-Informatics, Bio-Technology and Environmental Engineering - ABBE 2019

ISBN No. 978-1-63248-180-1 DOI : 10.15224/978-1-63248-180-1-03

Advanced diagnosis of Alzheimer's disease by automatically obtaining the best coronal slices for multi-classification recognition

Antonio Carrillo, Ignacio Rojas, Olga Valenzuela

Abstract. **The goal of this contribution is to find out a set of Y slices (coronal slices) from MRIs of patients with Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI), and Normal images, that provides the maximum accuracy in a multiclass classification system. Images are preprocessed and 2D wavelet coefficients are extracted to form a feature matrix. Using a feature selection algorithm called mRMR, the best features from the matrix are extracted; then, the dimension of the feature vectors is reduced using PCA and finally, it is used to train an SVM to perform multi-class classification. In order to find the best combinations of coronal slices, a multi-objective genetic optimization methodology based on NSGA-II is used and a set of different solutions are extracted from the Pareto front. More relevant solutions are selected using more flexible criteria than that of the Pareto front, and examine what slices and accuracies are achieved. The multi-classification accuracies obtained by the proposed method are 94.2% (the individual of the Pareto front with the highest accuracy) or 91.9% (using the best 13 slices according to their frequency of presence in the Pareto front). It is important to note that not only a good accuracy is obtained in the classification, but also new knowledge about the most relevant coronal slices to distinguish the four MRI classes (AD, Normal, MCI, LMCI)..**

***Keywords*- Alzheimer's Disease (AD), Magnetic Resonance Image (MRI), Discrete Wavelet Transform (DWT), minimum Redundancy Maximum Relevance (mRMR), Support Vector Machine (SVM), Genetic Algorithms (GA).**

I. INTRODUCTION

There are several possible causes for dementia, but Alzheimer's Disease (AD) is leading cause of dementia in the world. There are an estimated 35.6 million people in the world with dementia and more than 18 million of them have Alzheimer's disease, which represents more than 50% of the total people with dementia. Besides, the number of people with dementia is expected to increase to 65.7 million in 2030 and 115.4 million in 2050. Today, the diagnosis of Alzheimer's disease is made by using clinical criteria; however these criteria are not capable of diagnosing the disease in its pre-clinical stage, not allowing for an early diagnosis. Today, the diagnosis of Alzheimer's disease is made by using clinical criteria; however, these criteria are not able to diagnose the disease in its preclinical stage, not offering an early diagnosis.

Antonio Carrillo, Ignacio Rojas and Olga Valenzuela
University of Granada
CITIC-UGR, Granada (Spain)

Given this situation, it is necessary to develop methods and techniques to be included in the criteria that provide an early diagnosis, which would allow people with dementia to plan ahead while they still have the capacity to make important decisions about their future care as well as it would allow them to access available drug and non-drug therapies that may improve their cognition and enhance their quality of life.

At this point, it is necessary to identify the condition prior to dementia which is Mild Cognitive Impairment (MCI). Subjects with MCI are in an intermediate clinical situation between normality and dementia, which is characterized by the presence of subjective cognitive complaints but that do not make a significant alteration in activities of daily living. As there are studies that show that between the 10% and 15% of patients with MCI have developed dementia within a year, it is really important to be able of identify this pathology. Thus, it is in this field where imaging techniques can play a key role in early detection of patients who can develop dementia, in differential diagnosis of distinct dementias and monitoring the progression of the disease.

It is important to note that neuroimaging, using mainly magnetic resonance, is a powerful tool that adds a positive predictive value to the diagnosis and includes measurements using structural MRI to evaluate medial temporal lobe atrophy and positron emission tomography using fluorodeoxyglucose (FDG) or amyloid markers [1]. Many studies have focused on quantifying focal atrophy in the temporal lobe [2] and [3] and there are even visual scales to quantify the degree of atrophy, which are quick and easy to use. Recently, validations of computerized methods to measure the degree of temporal atrophy have been published. In comparison, these methods have a similar discriminatory power [4] with the advantage that they would facilitate measurements and would provide more objective results by standardizing the methods of analysis [5].

There are new developments in automatic classification systems based on computer intelligent paradigms, such as support vector machine (SVM), that present new diagnosis tools based on T1-weighted MRI [6],[7] and [8]. These approaches can be divided into three different categories, taking into account the type of features extracted from the MRI (voxel-based, vertex-based, or ROI-based) [1]. For the voxel-based category, the features are the probability of the

ISBN No. 978-1-63248-180-1 DOI : 10.15224/978-1-63248-180-1-03

different tissue classes in a given voxel [9]. In the second category, the vertex-based one, the features are defined at the vertex-level on the cortical surface. The methods based on ROI include mainly the analysis of the hippocampus (study of the volume, shape and specific characteristic of the hippocampus) .

Finally, it is important to mention some contributions that analyze in conjunction different sources of information databases. In the contribution presented by [11] the main goal is to analyze classification accuracies for detection of AD using FDG-PET, MRI, or combined information from both imaging modalities in two independent set of patients (ADNI database and Leipzig cohort). In this study, the authors arbitrarily extracted FDG-PET and MRI data of 28 AD patients and 28 healthy control subjects from the database provided by the Alzheimer's Disease Neuroimaging Initiative (ADNI) and compared them to data of 21 patients with AD and 13 control subjects from the Leipzig cohort. The authors have obtained high accuracies for discrimination between AD patients and control subjects in both cohorts using only FDG-PET or combined information from both imaging modalities. An interesting result is that using only MRI, the accuracies were substantially lower.

Even though these approaches achieve high accuracy (over 85\%), they were calculated on different studied populations (in fact, the variability between evaluations statistically increases mainly in case that the number of subjects is low; typically, there are several approaches in which less than a hundred subjects are used) making it difficult to compare the obtained results. It is important to note that most of articles published to date using intelligent classification systems in Alzheimer's disease diagnosis have three major drawbacks:

- They use a low number of data for both training and test (in some cases, less than a hundred).
- Classification is performed between patients with dementia and healthy patients, but it does not make any difference between patients with Mild Cognitive Impairment (MCI), Middle Cognitive Impairment Converters (MCI-C), and Middle Cognitive Impairment Non Converters (MCI-NC), since it is a complex task.
- They use only one type of feature (voxel-based, vertex-based or ROI-based), but not all together.

II. MATERIALS

A. Subjects cohort

The diagnostic classification was managed using a set of 255 subjects from the ADNI database and grouping them as AD (Alzheimer's disease), MCI (Mild cognitive impairment), Late Mild Cognitive Impairment (LMCI), and Normal.

The AD group contained 68 subjects ranging in age from 62 to 89 (74.3±7.6) years. The Mini Mental State Examination (MMSE) and Clinical Dementia Ratio (CDR) scores are mean 23.08 ±2.49 and mean 0.72 ±0.30, respectively. It is important to remember that a CDR of zero represents no dementia, and a CDR of 0.5, 1, and 2 indicate very mild, mild and moderate dementia respectively. The normal group has a total of 69 subjects, MCI 79, and finally, LMCI 49.

B. Segmentation and normalization

The images of the subjects are first segmented using the Segmentation routine implemented in SPM12. Afterwards, they are normalized to MNI space, using 1x1x1 mm voxels and a Bounding Box of limits: [-78 -112 -60; 78 76 85]. Since several articles report that the use of Gray matter achieves better results than the use of White matter or the whole brain, and in order to reduce the continue the experiment using just that Gray matter.

C. Feature extraction, selection, and reduction

Prior to feature extraction, slices used are intensity-normalized and filtered with a Gaussian filter with sigma=0.5. In order to extract the features from the MRI images, wavelet analysis will be used. In wavelet analysis, a fully scalable modulated window is shifted along the signal for every position, and this process is repeated with shorter or longer windows, yielding a multi-resolution analysis. As a result of this process when applied to an image, decomposing the original image in its wavelet coefficients (calculated as explained below) will produce a series of images with different scales.

There is a consensus among the literature regarding this topic that Db-4 wavelet function up to level 2 has the best trade-off between efficiency and computational cost, so the 2D Wavelet coefficients are obtained using Debuchie function of order 4 up to level 2 for each slice. Feature selection is achieved using the minimum Redundancy Maximum Relevance algorithm by [14].

D. Classification

For classification purposes, a Support Vector Machine (SVM) is used. In order to train and test an SVM model, it is first trained with a training set for which features and target labels are known. Hyperplane parameters are fit so that the model achieves its best accuracy separating the different data points, and then a testing set for which labels are known is fed into the SVM model and classified, getting a vector of labels as output. This output labels are compared to the actual target labels of the input feature vectors and the SVM model accuracy is obtained. Mathematically, the construction of the hyperplanes and classification.

E. Optimization

In order to find the best set of coronal slices to perform the classification, a Multi-Objective Genetic Algorithm (MOGA), creating a population of individuals consisting on bit strings of 161 bits, each bit coding whether a specific

Proc. of the Ninth Intl. Conf. on Advances in Bio-Informatics, Bio-Technology and Environmental Engineering - ABBE 2019

ISBN No. 978-1-63248-180-1 DOI : 10.15224/978-1-63248-180-1-03

slice should be selected or not. In this contribution, slices from Y=20 to Y=180 are only tested. The two fitness functions that need to be optimized by the algorithm (minimized) are: f(1)=100-accuracy(%) and f(2)=number_slices.

A population of 100 individuals is used, and the optimization is run for 200 generations. Each individual of each generation was stored in a data-file along with the accuracy achieved, so that a post-optimization analysis could be performed.

III. RESULTS

The most usual representation of the results obtained by a GA algorithm is that of a pareto front, shown in Figure 1. This graph shows the dominant individuals.

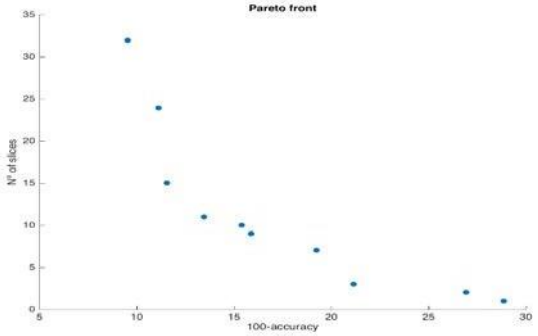

Figure 1 Pareto front. This graph shows the trade-off between accuracy and number of slices used

For the experiments carried out in this contribution, the minimum slices required for each achievable accuracy (from the set of accuracies achieved by all individuals) is also computed, as shown by Figure 2, and later, the best accuracy achieved by each number of minimum slices is computed. In a way, this is a more ``relaxed" version of the pareto front, and from now on we will refer to it as ``relaxed pareto".

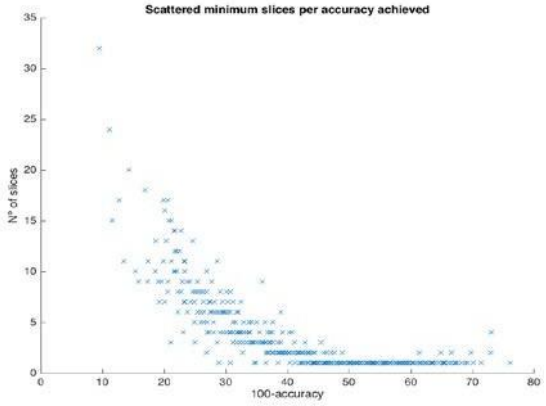

Figure 2 Minimum slices per accuracy achieved. It can be seen that for each number of slices, different accuracies can be achieved. The idea of the relaxed pareto came from selecting the highest accuracy achieved for each number of slices selected.

The histograms of slices for both sets of solutions have been obtained, as shown in Figures 3 and 4. To sum up, two sets of solutions are obtained: those of the pareto front, and those of the relaxed pareto.

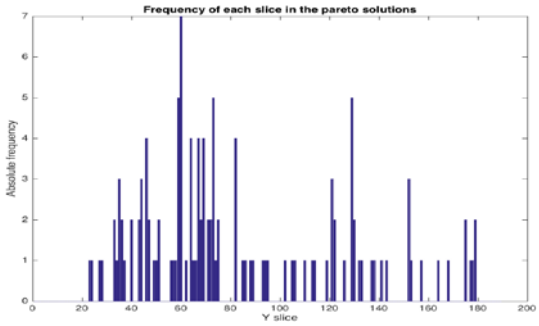

Figure 3 Histogram for the pareto front. There are several dominant slices: 60, 59, 73, 129....

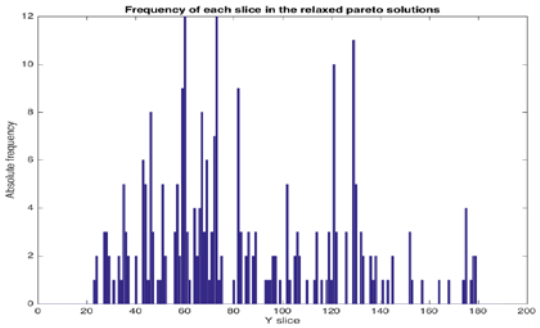

Figure 4 Histogram for the relaxed pareto. Its similarity to the pareto front supports our thinking that it could be a more flexible and extended version of the pareto front.

Observing the histograms, it is possible to get an idea of the set of best coronal slices to classify subjects. Since the genes of the individuals that are outperformed by others tend to disappear with further generations, the dominant individuals, which have dominant genes, exhibit a similar
genome, so that the frequency of each slice in the pareto front is somewhat an approximation of the relevance of that slice when extracting features for classification.

Table 1 AD, MCI, LMCI, and Normal subjects from the cohort

Frequency	Slices
7	60
5	59, 73, 129
4	46, 64, 67, 69, 82
3	35, 44, 121, 152
2	33, 36, 40, 43, 47, 51, 68, 71, 72, 75, 122, 130, 175, 179
1	23, 24, 27, 28, 34, 37, 49, 50, 56, 57, 58, 62, 65, 66, 74, 85, 86, 88, 89, 93, 94, 95, 102, 105, 106, 110, 113, 114, 119, 126, 132, 133, 137, 138, 141, 143, 153, 157, 164, 168, 177, 178

Table 2 Frequency of slices in the pareto front

Proc. of the Ninth Intl. Conf. on Advances in Bio-Informatics, Bio-Technology and Environmental Engineering - ABBE 2019

ISBN No. 978-1-63248-180-1 DOI : 10.15224/978-1-63248-180-1-03

Frequency	Slices
12	60, 73
11	129
10	121
9	59, 82
8	46, 67
7	72
6	43, 69
5	35, 44, 51, 57, 102, 130
4	64, 66, 175
3	27, 28, 36, 47, 56, 61, 68, 71, 83, 86, 89, 106, 114, 119, 122, 126, 132, 152
2	24, 29, 33, 37, 40, 52, 58, 65, 75, 85, 88, 96, 97, 105, 107, 133, 136, 138, 145, 178, 179
1	23, 31, 34, 45, 49, 50, 62, 70, 74, 80, 93, 94, 95, 99, 103, 110, 113, 116, 118, 120, 137, 141, 143, 153, 157, 164, 168, 174, 177

The ranking of slices obtained from the pareto front is shown in Table 1. Those of the relaxed pareto, in Table 2.The ``genome", consisting on the set of slices coded by each individual, both of the pareto front and relaxed pareto is shown in Figures 5 and 6.

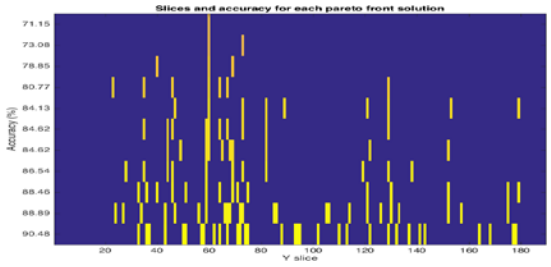

Figure 5 “Gene chart" of pareto front

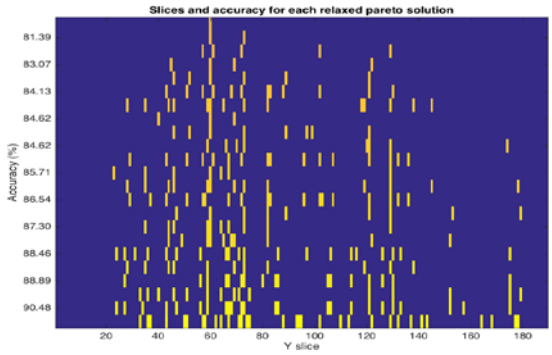

Figure 6 ”Gene chart" of relaxed pareto

In order to contrast the obtained results in this contribution, comparisons with those of others experts and researches presented in the bibliography have been carried out. Figures 7 to 10 clearly show consistency among different studies and researches. Specifically, the slices from the pareto front that appear on at least 3 of the individuals conforming the pareto have been selected, which happens to be 13 slices (sorted by relevance: 60, 129, 73, 59, 82, 69, 67, 64, 46, 152, 121, 44, 35).

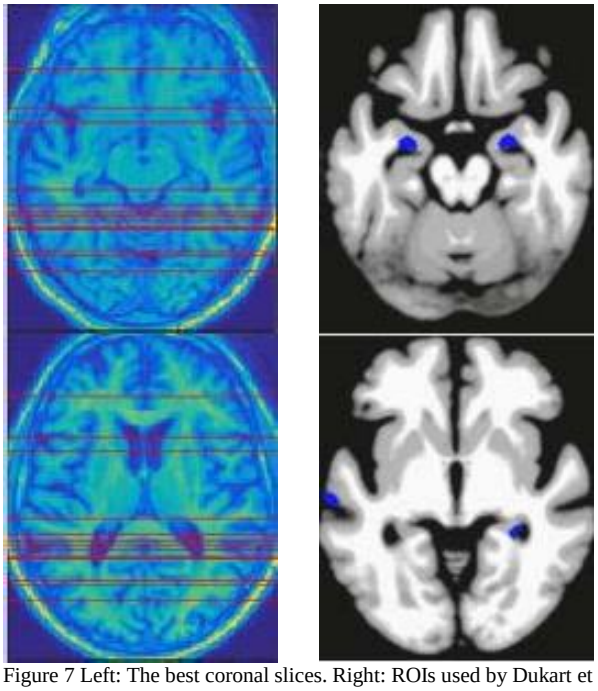

Figure 7 Left: The best coronal slices. Right: ROIs used by Dukart et al. 2012 [11]

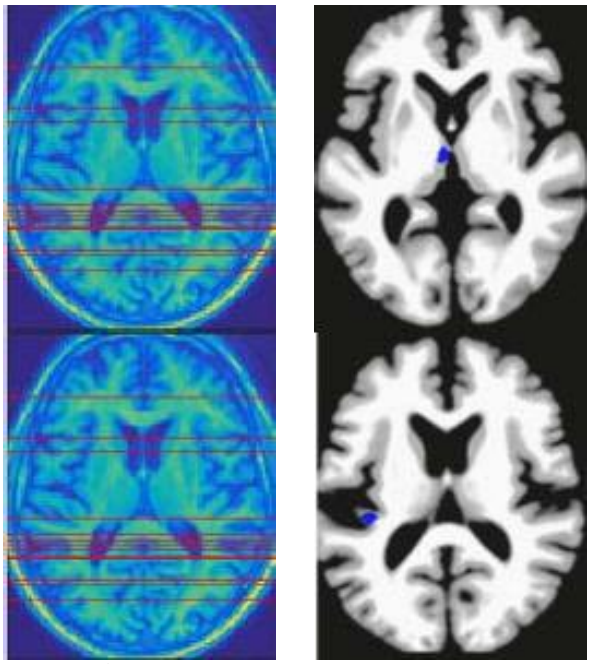

Figure 8 Left: The best coronal slices. Right: ROIs used by Dukart et al. 2012 [11]

Proc. of the Ninth Intl. Conf. on Advances in Bio-Informatics, Bio-Technology and Environmental Engineering - ABBE 2019

ISBN No. 978-1-63248-180-1 DOI : 10.15224/978-1-63248-180-1-03

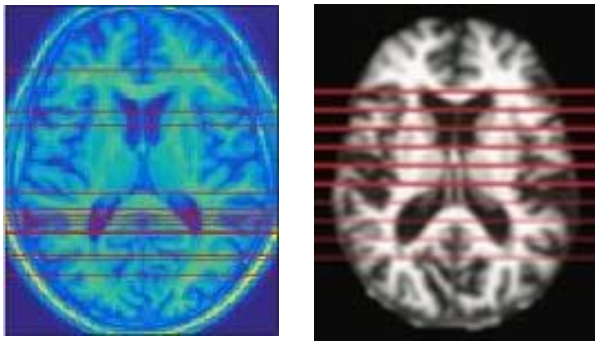

Figure 9 Left: The best coronal slices. Right: ROIs used by Zhang et al. 2015 [13]

Regarding the results given by Zhang et al. 2015 [13], it is important to highlight the fact that our histogram of best slices found by the genetic algorithm supports the results obtained by Zhang et al. and what they called Inter-Class Variance. Also, Zhang et al. 2015 showed that the slice Y=60 was the one that showed the highest relevancy when detecting AD malformations. Hence, slice Y=60 being the most relevant in the pareto front is a proof that the algorithm found the most relevant slices.

Because of the nature of the proposed method, which is that of a multi-class classifier, in order to test the classification accuracy of the system, confusion matrices are used.

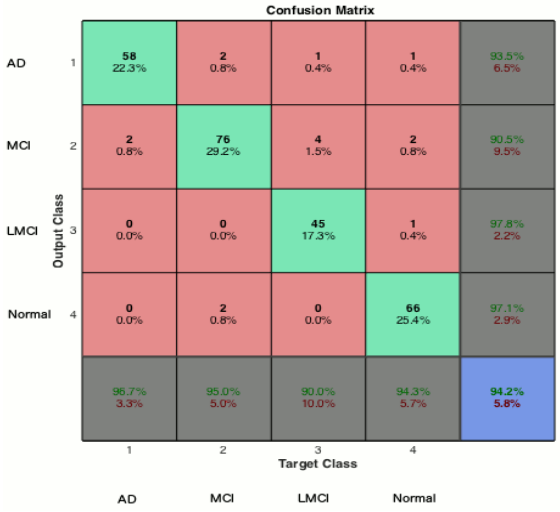

Figure 10 Confusion matrix, individual in the pareto front with the highest accuracy

A confusion matrix is a graphic representation of the performance of a multi-class problem where rows represent the output class given by the classifier and columns represent the target class, ie. the actual class. For our system, the confusion matrix is presented in Figure 10.

IV. CONCLUSION

In this paper, a new procedure to find the most significant ROIs when determining mental state in the development of Alzheimer's Disease have described. In order to find the best set of coronal slices to perform the classification, a Multi-Objective Genetic Algorithm based on NSGA-II is used. In this way, the training time required and the necessary computing time in order to find the best ROIs is significant. Prior researches and papers show good methods which achieve high accuracies, but the literature lacked of high classification accuracy when multi-class classification is performed. In most researches, subjects are classified class vs class. It would be an interesting development to design a method in which, given an MRI input, the software determined to which class it belongs and which are the most relevant coronal slices for this purpose. This could help medical experts in the diagnosis of different states in the development of AD.

Usually, ROIs determined by medical experts are used as the source of features. This contribution is faced with these issues and tries to investigate this computationally, so that maybe-hidden ROIs which experts had not noticed could be found. From the experiment carried out in this report, it is possible to assert that the obtained ROIs in our experiment are consistent with the existing bibliography.

The proposed method (without the requirement of prior knowledge from expert) is scalable and reproducible to further analyze any other neurological disease that affects gray or white matter; hence, the same procedure would be suitable to find relevant ROIs in other diseases.

ACKNOWLEDGMENTS

The publication of this article was supported by project RTI2018-101674-B-I00.

REFERENCES

[1] K Rémi Cuingnet, Emilie Gerardin, Jérôme Tessieras, Guillaume Auzias, Stéphane Lehéricy, Marie Odile Habert, Marie Chupin, Habib Benali, and Olivier Colliot. Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. NeuroImage, 56(2):766–781, 2011.

[2] G B Frisoni, P h Scheltens, S Galluzzi, F M Nobili, N C Fox, P H Robert, H Soininen, L-O Wahlund, G Waldemar, and E Salmon. Neuroimaging tools to rate regional atrophy, subcortical cerebrovascular disease, and regional cerebral blood flow and metabolism: consensus paper of the EADC. Journal of neurology, neurosurgery, and psychiatry, 74(10):1371–81, 2003.

[3] EL Koedam, M Lehmann, WM van der Flier, P Scheltens, YA Pijnenburg, N Fox, F Barkhof, and Wattjes MP. Visual assessment of posterior atrophy development of a MRI rating scale. Eur Radiol., 21:2618–2625, 2011.

[4] Eric Westman, Lena Cavallin, J. Sebastian Muehlboeck, Yi Zhang, Patrizia Mecocci, Bruno Vellas, Magda Tsolaki, Iwona Koszewska, Hilkka Soininen, Christian Spenger, Simon Lovestone, Andrew Simmons, and Lars Olof Wahlund. Sensitivity and specificity of medial temporal lobe visual ratings and multivariate regional MRI classification in Alzheimer's disease. PLoS ONE, 6(7), 2011.

[5] CR Jr Jack, F Barkhof, MA Bernstein, M Cantillon, PE Cole, C Decarli, B Dubois, S Duchesne, NC Fox, GB Frisoni, H Hampel, DL Hill, K Johnson, JF Mangin, P Scheltens, AJ Schwarz, R Sperling, J Suhy, PM Thompson, M Weiner, and NL. Foster. Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. Alzheimers Dement., 4(7):474–485, 2011.

ISBN No. 978-1-63248-180-1 DOI : 10.15224/978-1-63248-180-1-03

[6] Rik Vandenberghe, Natalie Nelissen, Eric Salmon, and et al. Binary classification of F-18-flutemetamol PET using machine learning: Comparison with visual reads and structural MRI. Neuroimage, 64:517–525, 2013.

[7] I Beheshti and H Demirel. Probability distribution function-based classification of structural MRI for the detection of Alzheimer's disease. Computers in biology and medicine, 64:208–16, 2015

[8] Edward Challis, Peter Hurley, Laura Serra, Marco Bozzali, Seb Oliver, and Mara Cercignani. Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI. NeuroImage, 112:232–243, 2015..

[9] B. Magnin, L. Mesrob, S. Kinkingnéhun, M. Pélégrini-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehéricy, and H. Benali. Support vector machinebased classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology, 51(2):73–83, 2009.

[10] N. Aggarwal, B. Rana, and R.K. Agrawal. 3d discrete wavelet transform for computer aided diagnosis of Alzheimer's disease using t1-weighted brain MRI. International Journal of Imaging Systems and Technology, 25(2):179–190, 2015

[11] Juergen Dukart, Karsten Mueller, Henryk Barthel, Arno Villringer, Osama Sabri, and Matthias Leopold Schroeter. Meta-analysis based SVM classification enables accurate detection of Alzheimer's disease across different clinical centers using FDG-PET and MRI. Psychiatry Research - Neuroimaging, 212(3):230–236, 2013.

[12] Kalyanmoy Deb. Multi-Objective Optimization Using Evolutionary Algorithms, 2001.

[13] Yudong Zhang, Zhengchao Dong, Preetha Phillips, Shuihua Wang, Genlin Ji, Jiquan Yang, and Ti-Fei 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, 9(June):1–15, 2015

[14] H C Peng, F H Long, and C Ding. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(8):1226–1238,2005.