# A NOVEL FEATURE SELECTION IN THE CASE OF BRAIN PET IMAGE CLASSIFICATION

*Imène Garali [a,b,c], Mouloud Adel [a], Salah Bourennane [a], Eric Guedj [b,c]*

[a] Aix-Marseille Université, CNRS, Centrale Marseille, Institut Fresnel UMR 7249, 13013 Marseille, France.
[b] Institut de Neurosciences de la Timone UMR 7289, Aix Marseille Université, CNRS, 13385 cedex 5, Marseille, France
[c] Centre Européen de Recherche en Imagerie Médicale, Faculté de Médecine, Marseille Aix-Marseille Université, France.

## ABSTRACT

Positron Emission Tomography (PET) imaging is of importance for diagnosing neurodegenerative diseases like Alzheimer Disease (AD). Computer aided diagnosis methods could process and analyze quantitatively these images, in order to better characterize and extract meaningful information for medical diagnosis. This paper presents a novel computer-aided diagnosis technique for brain PET images classification in the case of AD. Brain images are first segmented into Regions Of Interest (ROI) using an atlas. Computing some statistical parameters on these regions, we define a Separation Power Factor (SPF) associated to each region. This factor quantifies the ability of each region to separate AD from Healthy Control (HC) brain images. Ranking selected regions according to their SPF and inputting them to a Support Vector Machine (SVM) classifier, yields better classification accuracy rate than when inputting the same number of ranked regions extracted from four others classical feature selection methods.

***Index Terms***— Computer-Aided diagnosis (CAD), Support Vector Machine (SVM), Voxel-Based Analysis (VBA), Classification, Receiver Operating Characteristic (ROC), Alzheimer's Disease (AD), Feature Selection.

## 1. INTRODUCTION

Alzheimer Disease (AD) affects more than 30 million people over the world constituting most common cause of dementia. Positron Emission Tomography is a functional imaging modality which helps to diagnose such a disease. Early detection and diagnosis of AD is necessary for developing new treatments. Visual evaluation of brain scan images remains qualitative and operator-dependent. Computer-Aided Diagnosis (CAD), based on medical imaging, may be a valuable tool and could bring a quantitative evaluation to better detect and evaluate brain diseases such as AD.

Computer-aided diagnosis methods have been applied to tackle AD diagnosis [1-4]. They can roughly be divided into two groups: pixel and region based approaches. Pixel based techniques consider an image as a set of voxels, than compute some statistical parameters [5-9] and input them to a classifier. The main disadvantage of these methods is the small size problem where the number of voxels is higher than the images of the database. Region based techniques try to select ROI either using some knowledge on the pathology disorder [10-15] or some pattern analysis methods [16-19].

In this paper, we are concerned with metabolic activity in the gray matter rather than the entire brain since AD is a cortical disease. These intensity values for HC and AD images will allow us to classify the images accurately. Even when extracting only these voxels, there will still remain a very large number of voxels to input as attributes to a classifier, which is the main disadvantage of the Voxel-Based Analysis (VBA). Feature selection and reduction is the way within we can describe most but not all of the variance within your data, there by retaining the relevant information, while reducing the amount of redundant information to use. In this paper is a novel feature selection process consisting on an evolutionary search in the combinatorial space of features is presented. For that purpose, a novel region-based approach is developed to classify brain 18FDG-PET images. The motivation of this work is to identify the best regional features for separating AD patients from healthy controls, in order to reduce the number of features required to achieve an acceptable classification result while reducing computational time for the classification task.

First a brain mapping into regions given by an atlas is used. Each region is then characterized by a set of features computed on the moments of its grey-level histogram, to obtain a reduced feature space. Second, a ranking based on the ability of each region to best separate healthy from AD brain images is studied using Receiver Operating Characteristics (ROC) curves. Region selection results using the proposed method are compared to those obtained using classical feature selection techniques. When the best ranked regions for each feature selection method are input to the SVM classifier, the ones obtained by our approach outperformed classification results. The remain-

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

der of this paper is organized as follows. Section 2 describes PET image database and pre-processing. Section 3 presents the developed feature selection method and a short review of some state-of-the-art methods in the literature. Section 4 presents classification results. Finally, Section 5 gives our conclusions.

## 2. PET DATABASE

FDG PET scans were collected from the "La Timone" University Hospital, in the Nuclear Medicine Department (Marseille, France). The database was built up based on imaging studies of subjects that followed the standardized protocol of a hospital-based service. Therefore, FDG PET data was separated into two different classes: 61 healthy control (HC) (age range: 61.18[50…86] (mean [min…max])) and 81 patients with Alzheimer's disease (AD) (age range: 70.60[50…90] (mean [min…max])). HC were free from neurological/ psychiatric disease and cognitive complaints, and had a normal brain MRI. Patients exhibited NINCDS-ADRDA [20] clinical criteria for probable AD.

Since each patient has a different global metabolic rate, and the scaling of the images values depends on the specific scanner used, the data needed to be normalized. However, the proposed system will assume the same location within different images corresponds to the same anatomical position and the same grey-level intensity corresponds to the same brain activity, (among subjects).

The data were spatially normalized using SPM8 software [5]. The dimensions of the resulting voxel were 2x2x2 mm. Images were then smoothed with a Gaussian filter (8 mm FWHM) to blur individual variations and to increase signal-to-noise ratio [21].

After spatial normalization, intensity normalization was required in order to perform direct image comparison between different subjects. Intensity normalization was necessary since global brain activity varies from one subject to another. Different normalization methods were tested to determine the most appropriate one and were studied in [22]. These methods consisted in dividing the activity of voxels by a reference region activity, supposed to be preserved in patients. The gray-matter global brain ROI was used for that purpose.

## 3. FEATURE SELECTION AND REDUCTION

### 3.1. Combination Matrix

The brain was segmented into 116 Regions Of Interest (ROI) according to AAL of WFU-Pickatlas tool, version 2.4. [23]. For each ROI ($R_j$), the parameters mean ($\mu_1$), variance ($\mu_2$), skewness ($\mu_3$), kurtosis ($\mu_4$) and Entropy ($\mu_5$) are computed based on the region histogram h(x) as depicted in the following equations:

$$h(x) = \frac{number\ of\ pixels\ with\ gray\ level\ x}{total\ number\ of\ pixels\ in\ the\ region} \quad x \in \{0,..l\text{-}1\} \quad (1)$$

where $l$ is the number of the gray levels of a region (ROI) and $x$ is a gray level supposed to belong to {0, … $l$-1}.

$$\mu_1 = \sum_{x=0}^{l} x h(x) \quad (2)$$

$$\mu_i = \sum_{x=0}^{l} (x-m)^i h(x) \quad i = \{2, 3, 4\} \quad (3)$$

$$\mu_5 = -\sum_{x=0}^{l} h(x) \log_2(h(x)) \quad (4)$$

Our approach in this work consists on selecting ROIs which allow the best separation between our classes (AD and HC). Different parameter combinations for each region are used to select and rank ROIs according to a "Separation Power Factor" (SPF), defined in the following and which consequently produces a hierarchy of the ability of ROIs to separate between groups of subjects. The top-ranked ROIs are then introduced into a SVM classifier.

For each region, we examine the combination of these parameters (noted $p_1$, $p_2$, $p_3$, $p_4$ and $p_5$) with a length varying from 1 to 5. Therefore, we start by the combinations of length 1 ({$p_1$}… {$p_5$}), then those of length 2 ({$p_i$, $p_k$}, $1 \leq i,k \leq 5$, $i \neq k$), and so on until reaching the combination length of 5 parameters. Thereby, we create a combination matrix of $2^5$-1 = 31 columns. For each column, i.e. parameter combination, we compute the SPF (noted $\alpha_j$-$p_1p_{2...}p_i$) for each region $R_j$ $j \in$ {1...116} depends on the i parameters computed on it. The combination matrix has then 116 lines and 31 columns.

For a given region, HC and AD subjects are considered in an N-D feature space (N<=5). An N-Dimensional sphere (N-D sphere) is created over the group of healthy subjects (HC) (N=1 correspond to an interval, N=2 correspond to a disk, N>=3 correspond to a sphere).The N-D sphere's center is the mass center of a healthy subjects distribution. We depict in Fig.1 the case of 'Cingulum_Post_Left' ROI based on three parameters: the mean, the standard deviation and the kurtosis ($p_1$, $p_2$ and $p_4$).

At various radii, r, of the N-D sphere, we compute the following parameters:

- nb_HC_in, number of HC subjects inside the N-D sphere,
- nb_AD_in, number of subjects with AD inside the N-D sphere.
- nb_HC_in/nb_HC*, True Positive Rate (TPR).
- nb_AD_in/nb_AD*, False Positive Rate (FPR).

The ROC curve is created by plotting the fraction of true positives (i.e. HC subjects well ranked) vs. the fraction of false positives (i.e. the AD subjects misclassified) at various radii of the N-D sphere settings as it is presented in Fig.2. The Separation Power Factor (SPF) is defined as the area under ROC curve (AUC) and is within the range of [0, 1].

The set of the top ranked regions (associated to the higher values of SPF) are selected to be later used in classification experiments. The results of our study are shown in Table I where we considered the 3D brain image illustrated in Fig.3.

*nb_HC: number of HC subjects in the database
*nb_AD: number of AD subjects in the database

Throughout our work, physicians were regularly invited to discuss the appropriate value of the SPF threshold to be considered for region selection. Ultimately, the value of the threshold was set to 21ROIs. Indeed, previous simulations have shown that when the number of ROIs exceeds 21, a lower classification rate is achieved.

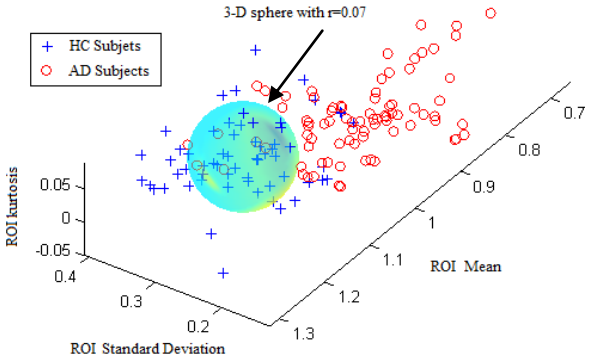


**Fig.1.** The separation between AD and HC groups relative to the region 'Cingulum_Post_Left' with three parameters: the mean, the standard deviation and the kurtosis.

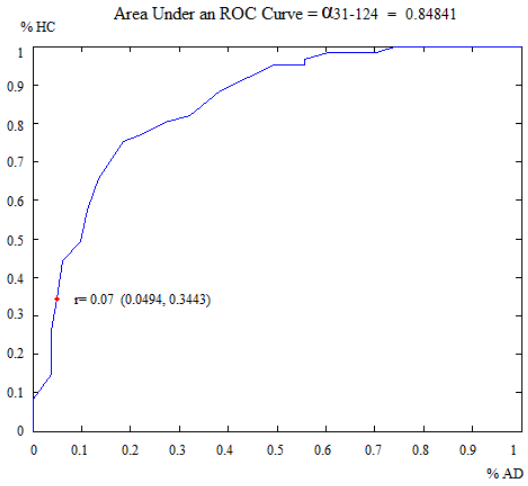


**Fig.2.** ROC curve obtained for region 'Cingulum_Post_Left' using three parameters: The mean, the Standard deviation and the kurtosis.

The obtained results, and in particular the 21ROIs identified are concordant with a recent review of the FDG-PET literature about the positive diagnosis of AD, with regions especially involving the temporo-parietal cortex, including the precuneus and the adjacent posterior cingulate cortex [24]. These regions will be considered in the early diagnosis of AD disease.

To compare our method to other feature selection methods, we present in the following some classical feature selection methods.

### 3.2. Fisher Score (FS)

Each ROI of PET image was characterized with five parameters. However, not all the ROIs have the same level of relevance in terms of discrimination between groups of

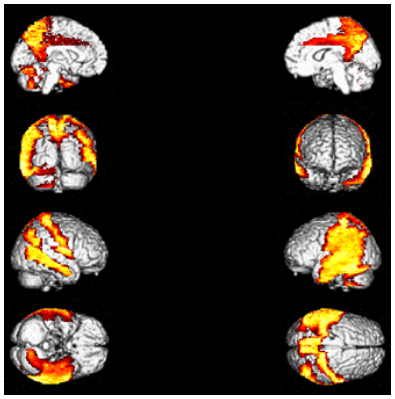

**Fig.3.** The 21 ROIs selected are presented on 3D brain image.

| Name of the 21ROIs | $\alpha_{j}$-$p_1p_2\ldots p_i$ | SPF |
|---|---|---|
| 'Cingulum_Post_L' | $\alpha_{31\text{-}124}$ | 0.8484 |
| 'Angular_R' | $\alpha_{4\text{-}13}$ | 0.8124 |
| 'Cingulum_Post_R ' | $\alpha_{32\text{-}12}$ | 0.8041 |
| 'Angular_L' | $\alpha_{3\text{-}12}$ | 0.8013 |
| 'Parietal_Inf_L' | $\alpha_{77\text{-}12}$ | 0.7958 |
| 'Temporal_Mid_L' | $\alpha_{99\text{-}13}$ | 0.7949 |
| 'Cingulum_Mid_R' | $\alpha_{30\text{-}3}$ | 0.7940 |
| 'SupraMarginal_L' | $\alpha_{95\text{-}12}$ | 0.7919 |
| 'Precuneus_L' | $\alpha_{85\text{-}1234}$ | 0.7907 |
| 'Rolandic_Oper_R' | $\alpha_{92\text{-}13}$ | 0.7861 |
| 'Temporal_Mid_R' | $\alpha_{100\text{-}13}$ | 0.7842 |
| 'Precuneus_R' | $\alpha_{86\text{-}123}$ | 0.7809 |
| 'Cerebelum_6_L' | $\alpha_{15\text{-}1234}$ | 0.7697 |
| 'Parietal_Inf_R' | $\alpha_{78\text{-}12}$ | 0.7615 |
| 'Cerebelum_6_R' | $\alpha_{16\text{-}12}$ | 0.7542 |
| 'Postcentral_R' | $\alpha_{82\text{-}1234}$ | 0.7542 |
| 'Temporal_Inf_L' | $\alpha_{97\text{-}12}$ | 0.7526 |
| 'Temporal_Sup_L' | $\alpha_{105\text{-}12}$ | 0.7476 |
| 'Fusiform_ L' | $\alpha_{53\text{-}1234}$ | 0.7471 |
| 'Cerebelum_8_L' | $\alpha_{19\text{-}1}$ | 0.7449 |
| 'Postcentral_L' | $\alpha_{81\text{-}123}$ | 0.7391 |

The SPF denoted $\alpha j\text{-}p_1p_2\ldots p_i$ is calculated for a region $R_j, j\epsilon$ {1...116} depending on the contribution of the parameters $p_i$ $i\epsilon$ {1.....5}.

**Table1.** The list of the 21 ROIs selected with the best parameters.

subjects. Thus, a regions selection process is done to retain only the most discriminant ones to differentiate between HC and AD patients. Thus, an initial regions selection based on discriminant ability is performed using the Fisher Score (FS) in order to obtain a vector of discriminants ROIs for our classes. The FS criterion is characterized by its separation ability for the two-class case. It may be defined as follows [25-26]:

$$FS_p = \frac{(\mu_{p1}-\mu_{p2})^2}{(\sigma_{p1})^2+(\sigma_{p2})^2} \quad (5)$$

where $\mu_{pi}$ and $(\sigma_{pi})^2$ denote the $i^{th}$ class mean value for the ROIs with the parameter $p$ varying from 1 to 5 and the variance for each input variable (ROIs with the parameter $p$), respectively. This way, the FS value increases as the ROI with the parameter $p$ is more discriminant between

the two classes. FS is used in this study to compute the most discriminating ROIs. Thereby, we create a matrix of 5 columns and 116 lines. The set of the top ranked regions (21 ROIs) with the higher value of FS are selected to be introduced into a SVM classifier.

### 3.3. SVM-RFE algorithm

Support Vector Machines Recursive Feature Elimination (SVM-RFE) algorithm proposed by Guyon [27] returns a ranking of the features of a classification problem by training a SVM with a linear kernel and removing the feature with the smallest ranking criterion. This criterion is the value of the decision hyperplane (denoted $\boldsymbol{w}$) given by the SVM. More detailed information, are given in [27]. In this analysis, the SVM-RFE input consists in 116 ROIs with five parameters. Thereby, we create a matrix of 116*5=580 columns and 116 lines.

### 3.4. Feature selection with Random Forests (RF)

Random Forest is an ensemble based classifiers that are often applied to high-dimensional dataset. In this work, random forest was applied to the feature selection. It is a classifier composed of many decision trees [25-26]. The number of trees grown in each forest, *t*, and the number of feature randomly selected at each tree node, *d*, had to be chosen. Several experiments were conducted to evaluate the random forest classifier before feature selection.
In this analysis, the random forest input consists in 116 ROIs with five parameters. Note that the generalization error converges to a limit as the number of trees in the forest becomes large. Moreover, the generalization error depends on the strength of the individual trees in the forest and the correlation between them. It can be concluded that, the random forest classifier converges for about 500-600 trees grown. We used t=595 and d=$\sqrt{D}$. At each node in a tree, d<<D features are randomly selected from the D available features in the dataset. As described in the random forest, estimates of relative importance of the features for classification may be extracted. These features are then injected into the SVM classifier.

### 3.5. Minimum Redundancy Maximum Relevance (mRMR)

It is common that a large number of features are not informative. The first idea of mRMR is that we should not use features which are highly correlated, i.e. the redundancy between features should be taken into account, thus keeping features which are maximally dissimilar to each other [28].

## 4. CLASSIFICATION RESULTS

Our study explored two methods: pixel and region based approaches for feature selection as well as how to use the parameters in each of these approaches: mono-parametric or multi-parametric analysis. In the case of Voxel-Based Analysis, we talk about pixel approach mono-parametric analysis (voxel is characterized by it intensity). In the case of others classical feature selection methods, we use a region based approach mono-parametric analysis. Each ROI was characterized with five parameters used independently. When, we selected the most discriminant ROIs to differentiate between HC and AD patients, each ROI is characterized by one parameter *p* varying from 1 to 5. In our contribution in this paper, a novel approach from region based technique multi-parametric analysis is presented. Each ROI depends on the i parameters computed on it, $1 \leq i \leq 5$, the main of our combination matrix method. This section provides the experimental results of evaluation of the CAD tool developed in this work. The classifier selected for this purpose case is a linear SVM. Table II provides the results of our approach using 21 ROIs with a combination matrix. In addition, the results of other techniques are provided for comparison.

### 4.1. Support Vector Machine (SVM)

As a general rule, a classifier takes a number of input variables qualifying the data under investigation and produces an output indicating the classes to which the data belong. Support vector machines implement a very simple idea. They map pattern vectors to a high-dimensional feature space where a 'best' separating hyperplane (the maximal margin hyperplane) is constructed. SVMs try to maximize the separation margin based on the distance between a hyperplane and the closest data samples. In the present work we used a SVM classifier with a linear kernel [25-26]. The performance of our classification approach is tested in three steps: training, cross-validation and test. Cross validation is achieved by means of the leave-one-out method, a technique that iteratively holds out a subject for test, while training the classifier with the remaining subjects, so that each subject is left out once.

| 21 ROIs for Feature Selection | SVM (Kernel = 'Linear') |
|---|---|
| ***Pixel approach mono-parametric analysis*** | |
| Voxel-Based Analysis | 94.36% |
| ***Region based approach mono-parametric analysis*** | |
| 116 ROIs with 5parameters | 92.95% |
| Score-Fisher | 88.73% |
| SVM-RFE | 85.91% |
| RF-Features selection | 91.54% |
| mRMR | 90.14% |
| ***Region based approach multi-parametric analysis*** | |
| **Combination Matrix** | **95.07%** |

**Table2.** The classification results (%) with different methods in the literature, for comparison.

When comparing the method presented in this work with other methods existing in the literature we can conclude that by injecting the 21ROIs with their combination of parameters in the SVM, we obtain the higher classification rate equal to 95.07% which is higher than that of the Voxel-Based Analysis equal to 94.36%, one of the main approaches which consists in studying the brain image as a set of raw voxels and input them to a classifier without any feature selection step.

## 5. CONCLUSION

In this work, a straightforward criterion to select a set of discriminant regions for the classification of PET brain images is presented. After normalization of the brain images, the set of the top ranked ROIs which presents greater overall difference between HC and AD brain images are selected. The combination parameters of the selected ROIs are used as features to the classifier. The use of the SPF shows better results than other methods when, we use the same regions provided by SPF method. In a future work, a combination parameter for ROIs selection for each method will be studied.

## 6. ACKNOWLEDGMENTS

This work was supported by the region PACA and Nicesoft.

## REFERENCES

[1] S. Minoshima, K. Frey, and R. Koeppe, "A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of Fluorine-18-FDG PET," *J. Nucl. Med.*, vol. 36, no. 7, pp.1238–1248, Jul. 1995.

[2] K. Ishii, A. Kono, H. Sasaki, N. Miyamoto, T. Fukuda, S. Sakamoto, and E. Mori, "Fully automatic diagnostic system for early- and lateonset mild Alzheimer's disease using FDG PET and 3d-SSP," *Eur. J. Nucl. Med. Mol. Imag*, vol. 33, pp. 575–583, 2006.

[3] G. Fung and J. Stoeckel, "SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information," *Knowledge Inf. Syst*., vol. 11, pp. 243–258, 2007.

[4] E. Westman, A. Simmons, Y. Zhang, J.-S. Muehlboeck, C. Tunnard, Y. Liu, L. Collins, A. Evans, P. Mecocci, B. Vellas, M. Tsolaki, I. Kloszewska, H. Soininen, S. Lovestone, C. Spenger, and L.-O. Wahlund, "Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls," *NeuroImage*, vol. 54, no. 2, pp. 1178–1187, 2011.

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

[6] D. Salas-González, J.M. Górriz, J. Ramírez, I.A. Illán, M. López, F. Segovia, R. Chaves, P. Padilla, C.G. Puntonet, "Feature selection using factor analysis for Alzheimer's diagnosis using F-FDG pet images", *Medical Physics* 37 (11) pp. 6084–6095, 2010.

[7] J. Stoeckel, N. Ayache, G. Malandain, P.M. Koulibaly, K.P. Ebmeier, J. Darcourt, "Automatic classification of SPECT images of Alzheimer's disease patients and control subjects", *Medical Image Computing and Computer-Assisted Intervention* – MICCAI, Vol. 3217 of Lecture Notes in Computer Science, Springer, Heidelberg, Germany, pp. 654–662, 2004.

[8] D. Salas-Gonzalez, J.M. Górriz, J. Ramírez, F. Segovia, R. Chaves, M. López, I.Á. Illán, P. Padilla, "Selecting regions of interest in SPECT images using Wilcoxon test for the diagnosis of Alzheimer's disease*", HAIS* (1), pp.446–451, 2010.

[9] W. Zhao, C. Wu, K. Yin, T.Y. Young, M.D. Ginsberg, "Pixel-based statistical analysis by a 3d clustering approach: application to autoradiographic images*", Computer Methods and Programs in Biomedicine* 83 (1), pp. 18–28, 2006.

[10] J. Ramírez, J. M. Górriz, D. Salas-González, A. Romero, M. López, I. Álvarez, M. Gómez-Río, "Computer-aided diagnosis of Alzheimer's type dementia combining support vector machines and discriminant set of features", *Information Sciences*, pp. 59–72, 2013.

[11] J. M. Górriz, A. Lassl, J. Ramírez, D. Salas-Gonzalez, C. Puntonet, and E. Lang, "Automatic selection of ROIs in functional imaging using Gaussian mixture models," *Neurosci. Lett,* vol. 460, no. 2, pp. 108–111, 2009.

[12] J. Stoeckel, G. Malandain, O. Migneco, P. Koulibaly, P. Robert, N. Ayache, and J. Darcourt, W. Niessen and M. Viergever, Eds., "Classification of SPECT images of normal subjects versus images of Alzheimer's disease patients," in *Med. Image Computing Computer-Assisted Intervention* (MICCAI 2001), Berlin, Germany, 2010, vol. 2208, pp. 666–674.

[13] M. López, J. Ramírez, J. M. Górriz, I. Alvarez, D. Salas-Gonzalez, F. Segovia, R. Chaves, P. Padilla, and M. Gómez-Río, "Principal component analysis-based techniques and supervised classification schemes for the early detection of Alzheimer's disease," in *Neurocomputing,* 2011, vol. 74, pp. 1260–1271.

[14] M. López, J. Ramírez, J. M. Górriz, I. Álvarez, D. Salas-Gonzalez, F. Segovia, and R. Chaves, "Automatic tool for the Alzheimer's disease diagnosis using PCA and Bayesian classification rules," *IET Electron. Lett*, vol. 45, no. 7, pp. 342–343, 2009.

[15] D. Salas-Gonzalez, J. M. Górriz, J. Ramírez, M. López, I. A. Illán, C. Puntonet, and M. Gómez-Río, "Analysis of SPECT brain images for the diagnosis of Alzheimer's disease using moments and support vector machines," *Neurosci. Lett*, vol. 461, no. 1, pp. 60–64, 2009.

[16] D. Zhang, D. Shen, "MultiCost: Multi-stage cost-sensitive classification ofAlzheimer's disease". *Machine learning in medical imaging.* Lect Notes ComputSci, 7009, pp. 344–51, 2011.

[17] W. Cai, D. Feng, R. Fulton. "Content-based retrieval of dynamic PET functionalimages", *IEEE Trans Inform Technol Biomed*, 4(2), pp.152–8, 2000.

[18] RA. Heckemann, S. Keihaninejad, P. Aljabar, D. Rueckert, JV. Hajnal, A. Hammers, "Improving inter-subject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation", *Neuroimage* 51(1), pp. 221–7, 2010.

[19] RA. Heckemann, S. Keihaninejad, P. Aljabar, KR. Gray, C. Nielsen, D. Rueckert, et al, "Automatic morphometry in Alzheimer's disease and mild cognitive impairment". *Neuroimage*, 56(4), pp. 2024–37, 2011.

[20] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, EM. Stadlan. "Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease". *Neurology*, Jul; 34(7), pp 939-44, 1984.

[21] J.A. Maldjian, P.J. Laurienti, R.A. Kraft, J.H. Burdette., "An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets", *NeuroImage* 19(3), pp. 1233-9, 2003.

[22] G.Imène, A. Mouloud, T. Sylvain, B. Salah, G. Eric, "Brain Region of Interest Selection for 18FDG Positrons Emission Tomography Computer-aided Image Classification.", *International Conference, IPTA 2014*, pp. 44-48

[23] R.P. Woods, *Spatial transformation models*, pp. 465-490, Chapter 29, Academic Press, San Diego, 2000.

[24] NI1. Bohnen, DS. Djang, K. Herholz, Y. Anzai, S. Minoshima. "Effectiveness and safety of 18F-FDG PET in the evaluation of dementia", A review of the recent literature. *J Nucl Med* 53(1), pp. 59-71, 2012.

[25] A.R. Webb, K.D. Copsey, *Statistical Pattern Recognition*, 3rd ed, pp.249-262, 389,439.

[26] Richard O.Duda, Peter E.Hart, David G. Stork., *Pattern Classification*, 2rd ed, pp.117-124, Chapter 3.

[27] I. Guyon, J.Weston, S. Barnhill, V. Vapnik, "Gene Selection for Cancer Classification using Support Vector Machines" *Machine Learning*, 46, pp. 389–422, 2002

[28] H.Peng, F.Long, 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*, Vol. 27, No. 8, pp.1226-1238, 2005.