Predicting Alzheimer's disease by classifying 3D-Brain MRI images using SVM and other well-defined classifiers

Article *in* Journal of Physics Conference Series · February 2012

DOI: 10.1088/1742-6596/341/1/012019

CITATIONS
8

READS
192

5 authors, including:

Sofia Matoug
Laurentian University
2 PUBLICATIONS **8** CITATIONS
SEE PROFILE

Amr R. Abdel-Dayem
Laurentian University
12 PUBLICATIONS **144** CITATIONS
SEE PROFILE

Kalpdrum Passi
Laurentian University
43 PUBLICATIONS **219** CITATIONS
SEE PROFILE

Mohammed Alqarni
McMaster University
4 PUBLICATIONS **11** CITATIONS
SEE PROFILE

Predicting Alzheimer's disease by classifying 3D-Brain MRI images using SVM and other well-defined classifiers

Predicting Alzheimer's disease by classifying 3D-Brain MRI images using SVM and other well-defined classifiers

S. Matoug*[1], A. Abdel-Dayem, K. Passi, W. Gross and M. Alqarni
Department of Mathematics and Computer Science
Laurentian University
935 Ramsey Lake Road Sudbury ON P3E 2C6 CANADA

E-mail: [1]sx_matoug@laurentian.ca

Abstract. Alzheimer's disease (AD) is the most common form of dementia affecting seniors age 65 and over. When AD is suspected, the diagnosis is usually confirmed with behavioural assessments and cognitive tests, often followed by a brain scan. Advanced medical imaging and pattern recognition techniques are good tools to create a learning database in the first step and to predict the class label of incoming data in order to assess the development of the disease, i.e., the conversion from prodromal stages (mild cognitive impairment) to Alzheimer's disease, which is the most critical brain disease for the senior population.

Advanced medical imaging such as the volumetric MRI can detect changes in the size of brain regions due to the loss of the brain tissues. Measuring regions that atrophy during the progress of Alzheimer's disease can help neurologists in detecting and staging the disease.

In the present investigation, we present a pseudo-automatic scheme that reads volumetric MRI, extracts the middle slices of the brain region, performs segmentation in order to detect the region of brain's ventricle, generates a feature vector that characterizes this region, creates an SQL database that contains the generated data, and finally classifies the images based on the extracted features. For our results, we have used the MRI data sets from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database[2].

1. Introduction

AD causes nerve cell death and tissue loss throughout the brain, resulting to brain tissue shrinking and larger ventricles (chambers within the brain that contain cerebrospinal fluid). When AD is suspected, the diagnosis is first confirmed with behavioural assessments and cognitive tests and often followed by

* To whom any correspondence should be made
[2] http://adni.loni.ucla.edu/

a brain scan [1]. Advanced medical imaging with computed tomography (CT) or magnetic resonance imaging (MRI), and with single photon emission computed tomography (SPECT) or positron emission tomography (PET) can be used to help exclude other cerebral pathology or subtypes of dementia [1]. Moreover, it may predict conversion from prodromal stages (mild cognitive impairment) to Alzheimer's disease [1].
The following report describes the whole process of pattern recognition where several steps are performed such as access ADNI database, describe the medical data, read the volumetric MRI, extract the middle slice of the brain region, perform segmentation methods in order to detect the region of brain's ventricle, generate a vector of attributes that characterizes this region, create an SQL database that contains the generated data, perform clustering to get the class labels and finally perform some classification methods based on the clustering results.

Section 2 describes briefly the organizational schema of the system implementation. Section 3 describes some of the tools used to access the medical ADNI database, gives an overview of the type of medical data files and discusses the problems encountered during the first step of accessing data. Section 4 includes the segmentation methods used to extract the regions of interest and shows the way they were used during the implementation process. Section 5 defines the vectors of attributes and shows their comprehensive statistical analysis. Section 6 introduces briefly the classification methods and shows the results of the used database in addition to classical databases results. Finally, in the conclusion, we summarize the different steps, assess the overall work and bring the future projection to the work.

2. Organizational schema of the system implementation

The proposed system is to perform 2D image segmentation of the original 3D MRI neuroimage brain data followed by attribute extraction and classification/prediction in order to assess, in the future, whether the patient is developing the AD and at which stage the disease stands. There are many existing segmentation techniques applied for medical image segmentation, including statistical methods, thresholding, edge detection, region-based techniques and more recently multi-resolution (using wavelets, ridgelets, etc.) techniques (figure 1, [18]). The choice of one method will depend on the type and quality of the image and the statistics of the extracted regions. The choice of attributes depends closely on the object's shape and statistics such as the statistical moments of order n, the textures, the geometric measures, etc. Once, we extract the attributes, we create a data base that is well defined and large enough to contain both training and testing data. A statistical analysis is made on the database to extract meaningful parameters and distributions that might assist the learning and classification processes. The first step of classification is to proceed to the learning process in order to produce the classes' categorical labels, and then we perform the classification. A step of accuracy measure of the classifier might be added to assess the accuracy of the employed classifier. The classification process is shown in figure 1.

3. Data access

The data we are using comes from Alzheimer's Disease Neuroimaging Initiative database (ADNI). ADNI is a multisite longitudinal clinical/imaging/genetic/biospecimen/biomarker study, whose goal is to determine the characteristics of AD as the pathology evolves from normal ageing to mild symptoms, to MCI, to dementia. It is a generally accessible data repository, which describes longitudinal changes in brain structure and metabolism.
ADNI uses several medical file formats such as the classical Analyse Format (hdr/img) that contains a header file and a separate 3*D* image and the next generation of medical images based on the Analyze Format, called NIFTI which is an *nii* structure containing both the header file and the 3*D* image.
The headers contain the information about the data such as the patient sex and age, the type of radiography, the view, size of voxels, etc, all of them stored into an info structure and the data itself as a 3D matrix usually of single type. All the nifti medical image files in ADNI database have the same standard which is: ADNI_pppp_S_ssss_Sequence_Sxxxx_Iyyyyy.nii where *pppp* is the patient ID,

ssss is the site ID, *Sequence* is the Sequence and processing steps, *Sxxxx* is LONIUID and *yyyyy* is the Image ID.

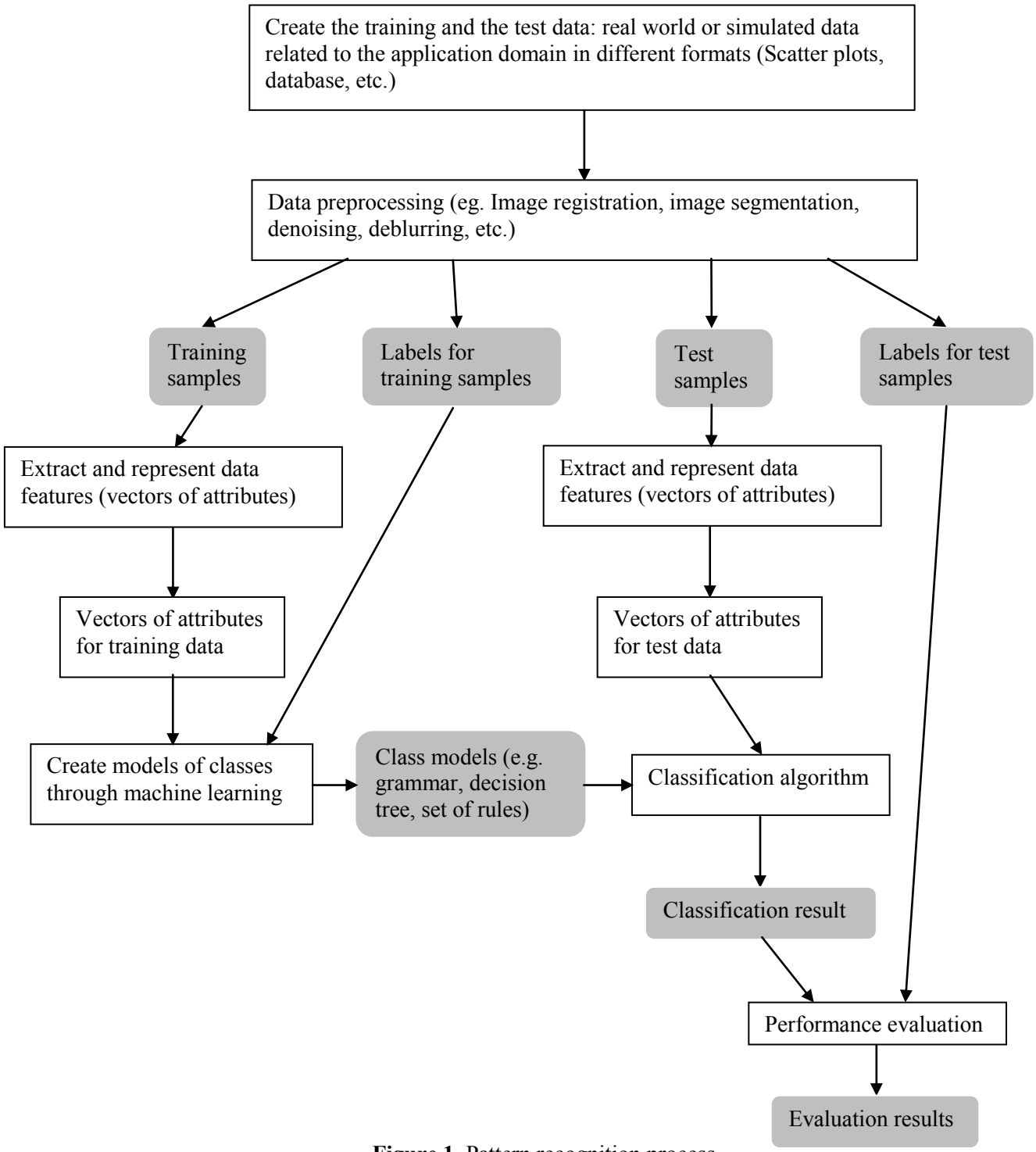

Figure 1. Pattern recognition process

In our work, we used Analyze and Nifti medical image formats without loss of generality, since both header files contain the same information even though the architecture of the structure is different

from one format to another. The data is first read using a matlab gui called readData3D[3], which allows the user to open medical 3D files. It supports the following formats : Dicom Files (.dcm , .dicom), V3D Philips Scanner (.v3d), GIPL Guys Image Processing Lab (.gipl), HDR/IMG Analyze (.hdr), ISI Files (.isi) and NifTi Files (.nii). etc.
Using the Matlab statement: [*D, info] = ReadData3D;*
Where D corresponds to the 3D MRI data and info is a structure containing the header information of the data such as the age, the sex, the view, etc. In case of Analyze 7.5 Format which uses radiological orientation (LAS), the data should be flipped for correct image display in MATLAB and reoriented for easier interpretation (stand patient up).
We also used other packages that allow 3*D* view and extraction of statistical information such as *Twfu_bpm* toolbox[4] for multimodal image analysis called biological parametric mapping (BPM), based on a voxel-wise use of the general linear model. It has a high degree of integration with the SPM[5] (statistical parametric mapping) software relying on it for visualization and statistical inference.
The medical data is accessed through the different packages and stored into *3D* arrays where each element corresponds to a voxel with the three space coordinates and the intensity value. Depending on the chosen view, we might flip or interchange the three dimensions of each medical data. The next step is to extract meaningful information from the data using segmentation methods on the middle slice by taking the middle 2D array (or image) from the transversal view of the *3D* data.

4. Segmentation

The segmentation is a well known tool that has the purpose of partitioning the image into regions of interest in order to extract 'meaningful' objects; several methods were developed to segment images based on thresholds, feature extraction and Clustering techniques.

4.1. Thresholding

When images contain different contrasting objects, thresholding provides effective means for obtaining segmented images. Thresholding techniques are based on partitioning the intensities using global or local threshold calculations techniques such as Otsu [3] and Niblack methods [12], where each threshold classifies the voxels in *3D* images (or pixels in 2D images) into different modes using a clustering criterion. Otsu method [13] is a clustering technique that tends to produce two tight clusters by minimizing their overlap (misclassified pixels). The threshold is adjusted dynamically by increasing the spread of one cluster and decreasing the spread of the other one. The goal then is to select the threshold that minimizes the combined spread.

4.2. Edge detection

Other segmentation methods are based on edge detection techniques such as Canny [4] , active contours or snakes using the technique of matching a deformable model to an image by means of energy minimization [5], [11]. The active contour, or snake algorithm was introduced by Kass [11]. Given an approximation of the boundary of an object in an image, an active contour model deforms the initial boundary to lock onto characteristic features within the region of interest. The contour is deformed iteratively until it matches the boundary of the region of interest by looking for the minimum of energy of a given problem. The energy functional is a weighted combination of internal and external forces depending on the shape of the snake and location within the image.

[3] http://www.mathworks.com/matlabcentral/fileexchange/29344
[4] http://www.fmri.wfubmc.edu/cms/software
[5] http://www.fil.ion.ucl.ac.uk/spm/

4.3. *Region-based*

Region-based segmentation uses different techniques such as seeded region-growing [2], split-and-merge [8], watershed [10] and Wavelet-based segmentation [9] which is based on mathematical concepts such as quadrature mirror filtering, sub-band coding and pyramidal image processing.
The accessed data, as explained in the previous subsection, shows slices of images that can be visualized into different views (coronal, transversal or longitudinal). Figure 2 shows a transversal view of the middle slice. Since the cost is very important, we noticed that some of the upfront or down-front slices, spatially low positions (below brain mass), show very low intensity values which leads to the choice of thresholding values that remove those slices in order de decrease the cost of the algorithms. Before segmenting the reduced slices, we noticed that the area of interest that corresponds to the most visible ventricle is located in the middle, which corresponds to the middle transversal slice. To do that, we used the Otsu global thresholding using custom GUI tool to select best threshold level values where the starting intensities correspond to the user chosen value. This will lead to ignore low level intensities (back-ground air, CSF and other soft tissues) which correspond to the intensity 21 in our case. Using the same Matlab gui, we ignored the high levels (skull and other hard Tissues) which correspond to the intensities above 1995. For the images presenting a big area around the ventricle chambers (figure 2), we used the active contour techniques to extract the ventricle chamber's region (figure 3). But for smaller areas, the latter algorithm was unsuccessful compared with the DRLSE segmentation which extracts the active contour using the Distance Regularized Level Set Evolution (DRLSE) formulation.

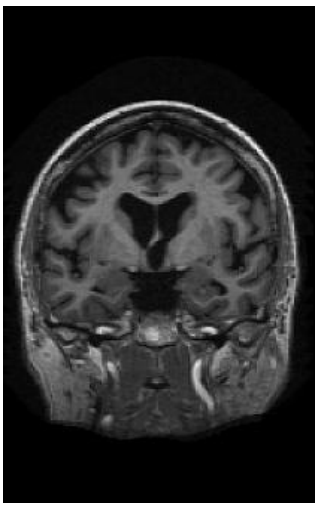

Figure 2. Transversal view of the middle slice

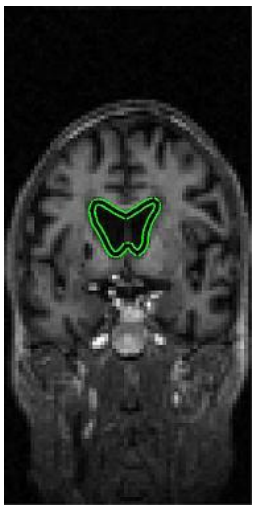

Figure 3. Active contour of a slightly reoriented and resized middle slice.

5. Attribute selection

We could assess the AD disease based on the shape of the ventricle chamber's area. Since AD causes the loss of brain mass due to molecules created in this area and being spread all over the brain, we assumed in this paper that we can assess the whole brain based on this specific area in order to decrease the cost of calculation. The more the chambers are big the more the AD is spread, which leads to the characterization of this area based on the shape of the chambers using statistical and geometrical attributes. The resulting attributes are respectively the surface area of the extracted region (Surf), the perimeter (Per), the first and second statistical moments (Mean, Std), 28 horizontal distances (W01, W02, …, W28), the height (Height) and the coordinates of the center of gravity of the region (Gx, Gy). The attributes are normalized into the range [0 1]. In order to make a first assessment

of the behavioral trend of each attribute, we made a statistical analysis of the 35 attributes using stacked columns where each attribute is summarized by a column bar containing five key data points (also called five-number summary) named max, third quartile (Q3), median, first quartile (Q1) and min. The median is the middle point of the classified set of data. The first and third quartile values are respectively the 25th and 75th percentiles. The first four attributes (surface, perimeter, first and second statistical moments) and the last three attributes (height and coordinates of the centre of gravity) tend to vary more clearly and the widths shown from the 17th o the 32nd bar carry very little information and tend to have similar values which clearly is not a good tool to classify between the different medical image data base (figure 4).

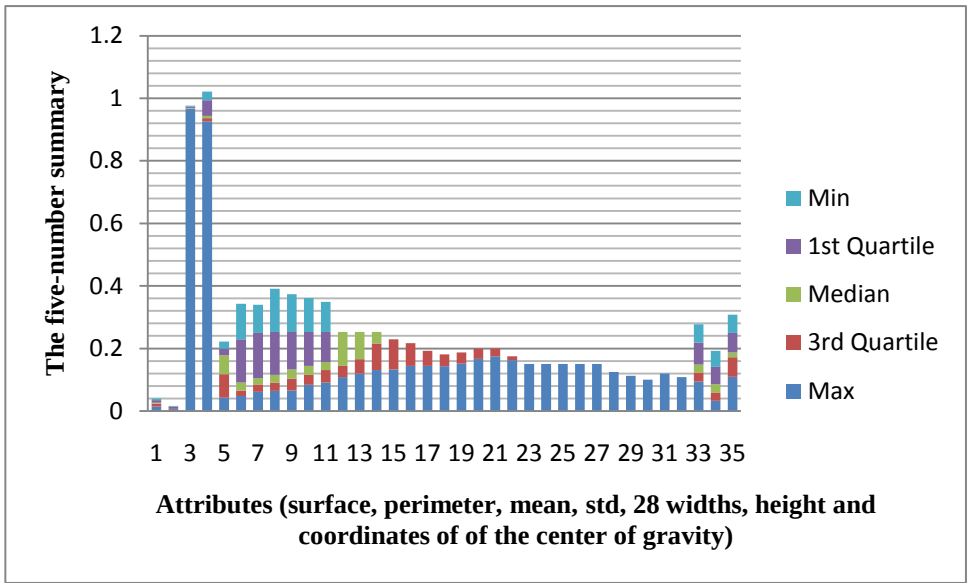

Figure 4. Statistical trend of 105 medical image data base based on the five-number summary (Min, 1st Quartile, Median, 3rd Quartile and Max)

6. Classification

The last step of the general framework is to proceed to the classification process starting by the K-Nearest Neighbors (KNN) clustering technique in order to label the classes and then creating a similarity criterion that classifies each data into the proper class name or label using the support vector machine (SVM) classification method. The SVM is a classifier that attempts to create a linear vector that segments the classes equally. In the event that is not linearly inseparable, the SVM algorithm lifts the space into a higher dimensional plane, until, on that plane, a vector can separate the classes. It is a particularly good classifier, albeit a slow one [6], [16], [17]. It supports various mathematical formulations. We used the C-Support Vector Classification or C-SVC [20], [21] that solves the following optimization problem

$$\begin{aligned} \boldsymbol{min}_{\boldsymbol{\omega},\boldsymbol{b},\boldsymbol{\xi}} \quad & \tfrac{1}{2}\boldsymbol{\omega}^{\boldsymbol{T}}\boldsymbol{\omega} + \boldsymbol{C}\textstyle\sum_{\boldsymbol{k}=\boldsymbol{0}}^{\boldsymbol{n}}\boldsymbol{\xi}_{\boldsymbol{i}} \\ \text{subject to} \quad & \boldsymbol{y}_{\boldsymbol{i}}\left(\boldsymbol{\omega}^{\boldsymbol{T}}\boldsymbol{\phi}(\boldsymbol{x}_{\boldsymbol{i}}) + \boldsymbol{b}\right) \geq \boldsymbol{1} - \boldsymbol{\xi}_{\boldsymbol{i}} \\ & \boldsymbol{\xi}_{\boldsymbol{i}} \geq \boldsymbol{0}, \boldsymbol{i} = \boldsymbol{1}, \dots, \boldsymbol{l}, \end{aligned} \qquad (1)$$

given training vectors $x_i \in R^n$, i=1, …,l, and an indicator vector $y \in R^l$
where $\phi(x_i)$ maps x_i into a higher-dimensional space and C > 0 is the regularization parameter.

Clustering is a procedure which aims to categorize objects. In particular, the objective is to categorize objects in groups that do not have class labels. Popular methods of clustering include; K-Nearest Neighbors, K-Means and K-Medoids [15], [7].
The results of clustering are summarized in tables 1 and 2 for the number of clusters k = 2, 5 and 7 respectively, showing the clusters' centroids of each attribute. Since our data contains 105 samples, If we choose k = 2 for two classes, based on the results of the KNN, we will get a vector of 105 labels of 1's and 2's for the 105 samples as shown in table 3.

Table 1: KNN results showing the clusters' centroids of each attribute for k=2 and 5

	k=2		**k=5**				
Attributes	**Class 1**	**Class 2**	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Class 5**
Surf	0.0098	0.0211	0.0059	0.0091	0.0219	0.0208	0.0043
Per	0.0053	0.0084	0.0039	0.0048	0.0081	0.0093	0.0024
Mean	0.0221	0.0118	0.0079	0.0089	0.0128	0.0094	0.9738
Std	0.0706	0.0626	0.0510	0.0613	0.0694	0.0535	0.9942
W01	0.0877	0.0853	0.1153	0.0679	0.0374	0.1346	0.0222
W02	0.1404	0.1708	0.1201	0.1446	0.1669	0.1723	0.0889
W03	0.1477	0.1926	0.1112	0.1614	0.2055	0.1736	0.0889
W04	0.1363	0.2023	0.0898	0.1542	0.2210	0.1715	0.0889
W05	0.1221	0.1999	0.0701	0.1431	0.2229	0.1612	0.0667
W06	0.1053	0.1920	0.0407	0.1298	0.2182	0.1505	0.0667
W07	0.0910	0.1820	0.0340	0.1124	0.2087	0.1403	0
W08	0.0589	0.1706	0.0053	0.0742	0.1965	0.1288	0
W09	0.0380	0.1576	0.0013	0.0453	0.1822	0.1144	0
W10	0.0196	0.1437	0.0013	0.0180	0.1693	0.0988	0
W11	0.0107	0.1260	0.0007	0.0073	0.1578	0.0769	0
W12	0.0073	0.1080	0	0.0057	0.1428	0.0570	0
W13	0.0038	0.0877	0	0.0005	0.1165	0.0469	0
W14	0.0035	0.0667	0	0	0.0852	0.0403	0
W15	0.0055	0.0539	0	0	0.0654	0.0421	0
W16	0.0082	0.0436	0	0	0.0538	0.0435	0
W17	0.0082	0.0381	0	0	0.0471	0.0408	0
W18	0.0077	0.0299	0	0	0.0376	0.0348	0
W19	0.0092	0.0212	0	0	0.0296	0.0329	0
W20	0.0098	0.0184	0	0	0.0262	0.0329	0
W21	0.0094	0.0181	0	0	0.0174	0.0375	0
W22	0.0078	0.0145	0	0	0.0054	0.0367	0
W23	0.0071	0.0159	0	0	0.0020	0.0390	0
W24	0.0065	0.0164	0	0	0.0007	0.0392	0
W25	0.0049	0.0164	0	0	0.0007	0.0348	0
W26	0.0046	0.0136	0	0	0.0013	0.0297	0
W27	0.0033	0.0111	0	0	0.0013	0.0229	0
W28	0.0030	0.0111	0	0	0.0013	0.0220	0
Height	0.1042	0.0762	0.0922	0.1147	0.0819	0.0749	0.0667
Gx	0.0903	0.0579	0.0813	0.0985	0.0577	0.0631	0.0667
Gy	0.0785	0.1393	0.0528	0.0705	0.1411	0.1489	0.0469

Table 2: KNN results showing the clusters' centroids of each attribute for k=7

Attributes	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Surf	0.0167	0.0043	0.0055	0.0228	0.0095	0.0081	0.0237
Per	0.0071	0.0024	0.0037	0.0083	0.0049	0.0047	0.0110
Mean	0.0124	0.9738	0.0088	0.0097	0.0089	0.0082	0.0094
Std	0.0624	0.9942	0.0520	0.0661	0.0623	0.0579	0.0498
W01	0.1468	0.0222	0.1082	0.0287	0.0178	0.1226	0.1197
W02	0.1871	0.0889	0.1140	0.1629	0.1327	0.1544	0.1584
W03	0.1928	0.0889	0.1066	0.2032	0.1598	0.1574	0.1572
W04	0.1921	0.0889	0.0825	0.2207	0.1536	0.1482	0.1540
W05	0.1856	0.0667	0.0597	0.2237	0.1417	0.1385	0.1400
W06	0.1735	0.0667	0.0285	0.2200	0.1266	0.1265	0.1297
W07	0.1620	0	0.0212	0.2116	0.1173	0.1014	0.1206
W08	0.1506	0	0.0063	0.1987	0.1093	0.0206	0.1094
W09	0.1385	0	0.0016	0.1843	0.0639	0.0116	0.0925
W10	0.1177	0	0.0016	0.1729	0.0229	0.0034	0.0792
W11	0.0840	0	0.0008	0.1646	0.0107	0.0034	0.0598
W12	0.0575	0	0	0.1517	0.0082	0.0029	0.0499
W13	0.0369	0	0	0.1297	0.0010	0	0.0474
W14	0.0249	0	0	0.0946	0	0	0.0497
W15	0.0214	0	0	0.0709	0	0	0.0596
W16	0.0149	0	0	0.0587	0	0	0.0698
W17	0.0114	0	0	0.0519	0	0	0.0680
W18	0.0049	0	0	0.0429	0	0	0.0625
W19	0.0028	0	0	0.0338	0	0	0.0616
W20	0.0021	0	0	0.0299	0	0	0.0624
W21	0	0	0	0.0203	0	0	0.0742
W22	0	0	0	0.0067	0	0	0.0725
W23	0	0	0	0.0036	0	0	0.0764
W24	0	0	0	0.0049	0	0	0.0733
W25	0	0	0	0.0070	0	0	0.0620
W26	0	0	0	0.0071	0	0	0.0526
W27	0	0	0	0.0064	0	0	0.0398
W28	0	0	0	0.0064	0	0	0.0380
Height	0.0626	0.0667	0.0903	0.0834	0.1053	0.1215	0.0982
Gx	0.0513	0.0667	0.0780	0.0585	0.0885	0.1095	0.0810
Gy	0.1315	0.0469	0.0512	0.1364	0.0778	0.0607	0.1654

Table 3. Class labels corresponding to the 105 samples (S1, S2, ..., S105) for k=2 (2 classes) using KNN clustering technique

S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**
1	2	2	2	1	1	2	2	2	2	2	2
S13	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
2	2	2	2	2	2	1	2	2	2	2	2
S25	**S26**	**S27**	**S28**	**S29**	**S30**	**S31**	**S32**	**S33**	**S34**	**S35**	**S36**
2	1	2	2	2	2	2	2	2	1	1	1
S37	**S38**	**S39**	**S40**	**S41**	**S42**	**S43**	**S44**	**S45**	**S46**	**S47**	**S48**
1	2	2	2	2	2	1	2	2	1	1	1
S49	**S50**	**S51**	**S52**	**S53**	**S54**	**S55**	**S56**	**S57**	**S58**	**S59**	**S60**
2	2	2	1	2	2	2	2	2	1	2	1
S61	**S62**	**S63**	**S64**	**S65**	**S66**	**S67**	**S68**	**S69**	**S70**	**S71**	**S72**
1	2	2	1	1	2	2	2	2	1	1	1
S73	**S74**	**S75**	**S76**	**S77**	**S78**	**S79**	**S80**	**S81**	**S82**	**S83**	**S84**
1	1	2	2	2	2	2	2	2	2	1	2
S85	**S86**	**S87**	**S88**	**S89**	**S90**	**S91**	**S92**	**S93**	**S94**	**S95**	**S96**
2	1	2	1	2	2	2	2	1	1	1	2
S97	**S98**	**S99**	**S100**	**S101**	**S102**	**S103**	**S104**	**S105**			
2	1	2	1	2	2	1	1	1			

For the SVM, we could not interpret the results since they represent 35 dimensions, however we can examine the results visually by plotting only 2-D vector of attributes. We set the parameter C of class i to *weight*C* in C-SVC in equation 1.

Using the first two attributes that correspond to the surface and perimeter of the objects and by forcing a linearly separable solution, we got a classification accuracy of 0% for 7 support vectors (SV). The weights are respectively -767.36739 and 1074.7484, the bias b_o = 4.6073 the margin = 0.0015145. The classification results are plotted in figure 5 showing a predominant *class 2* for the first two attributes.

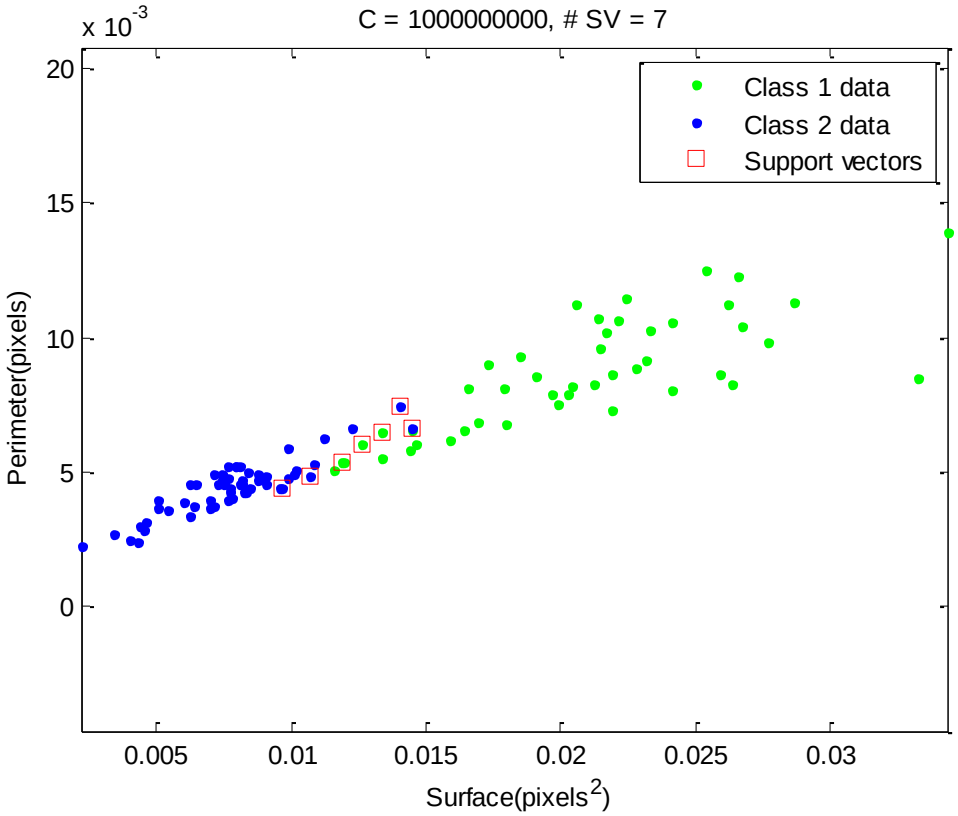

Figure 5. SVM classification results for attributes 1 and 2 (Perimeter vs. surface) with linearly separable solution

The same work was performed for other pairs of attributes. Table 4 summarizes the SVM classification results corresponding to 7 pairs of attributes by forcing a linearly separable solution. The results correspond to the number of support vectors, the weights, the parameter b_o, the margin and a visual classification assessment (good, fair, poor).

Table 4. SVM classification results for a pair of attributes with linearly separable solution

Attribute x	Attribute y	SV	weights	b_o	Margin	Visual assessment
Mean	Std	87	-6.8384 .8242	1.6761	0.2923	Poor
Width01	Width02	59	1.0811 6.4423	0.74054	0.30617	Poor
Width03	Width10	17	-6.7621 -17.4341	3.6638	0.10695	Good
Width10	Height	4	-14.6321 -2.5301	3.2781	0.13469	Good
Gx	Gy	36	7.5521 -6.5821	1.7852	0.19964	Fair
Surface	Height	14	95.2414 -7.4722	0.83633	0.020935	Good
Perimeter	Width11	13	131.9472 11.7253	0.19204	0.015098	Good

The results shown in table 4, imply a good classification in the case of geometric attributes (e.g the widths and the surface) and fair to poor classification in the case of statistical attributes (mean, standard deviation and center of gravity).
Other techniques of classification have been used but did not produce greater results due to the choice of the vector of attributes and the small size of the data base.

The first one is the Bayesian Network which is a decision making model. This model makes decisions based on the dependencies of the previous conditions using probability theory. It takes a set of particular inputs and based on the probability of the occurrence of the inputs in the training set, it makes a decision of where the new input is supposed to be classed. The highest probability calculated for a particular output based on the inputs is the decision that is finally chosen as the output [3], [6], [19].
The second method is the neural network which is a learning system that creates interconnecting artificial neurons. It receives an input through its input layer and produces an output that depends on the architecture of the neural network, the values of the initial weights and the number of hidden layers. The hidden layer contains a set of neurons that have various weights to produce and output from each neuron. Once each neuron has a result, those results are aggregated with another set of weights to produce the final output [3].

7. Conclusion

Medical image processing and machine learning tools can help neurologists in assessing whether a subject is developing the Alzheimer disease. A machine learning system has been developed in order to extract meaningful information from the ADNI database, where the ventricle chambers are extracted using a segmentation method based on the statistical features of the region of interest. As a first impression, we chose this region to see if it corresponds to a good marker, however, the shape and size of ventricle did not have a big impact on the final result and we noticed that some other regions might be more affected. New features should be generated in order to characterize the new region. Due to the time factor, we could not classify the regions using different methods as expected, however we created different tools and simulated the whole system to work for the real image database. The developed algorithms have been tested on real medical data from ADNI database.
What was successfully completed is the creation of a framework to continue researching the use of SVMs towards mining medical images. This framework has potential to make potent predictions based exclusively on the properties of a patient's hippocampus based on the strength of the SVM's ability to classify objects. With further research, a system could be developed, coupled with a better feature set to produce an early detection for patients.
Ongoing research is focusing on the enhancement of the classification algorithms, adding more data to the system, enhancing the segmentation tool and assessing the results of the classification.

References

[1] Schroeter ML, Stein T, Maslowski N, Neumann J 2009 Neural correlates of Alzheimer's disease and mild cognitive impairment A meta-analysis including 1351 patients. *NeuroImage* **47** 1196-1206.
[2] Adams, R., and Bischof, L. 1994 Seeded region growing. IEEE Transactions on Pattern Analysis and Machine Intelligence **16** 641-647.
[3] Bishop, C. M. 1995 Neural Networks for Pattern Recognition. (New York, Oxford University Press, Inc.).
[4] Canny, J.1986 A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI*-**8**, 6.
[5] Caselles, V., Kimmel, R., and Sapiro, G. 1997 Geodesic active contours. Int. J. Comput. Vision **22**, 61-69.
[6] Duda, R. O., Hart, P. E., and Stork, D. G. 2001 Pattern Classification, (New York, 2. ed. Wiley).
[7] Filippone, M., Camastra, F., Masulli, F., and Rovetta, S. 2008 A survey of kernel and spectral methods for clustering. Pattern Recogn. **41** 176-190.
[8] Horowitz, S., and Pavlidis, T. 1977 Picture segmentation by a directed split and merge procedure. CMetImAly77, 101-111.

[9] Mallat, S. G. 1989 A theory for multi-resolution signal decomposition: The wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell. **11** 674-693.
[10] Meyer, F. 2001 An overview of morphological segmentation. IJPRAI **15** 1089-1118.
[11] Kass M, Witkin A and Terzopoulos D, 1988 Snakes: Active contour models, Int. J. Comp. Vision **1** 321-331.
[12] Niblack, W. 1986 An Introduction to Digital Image Processing. (New Jersey, Prentice Hall)
[13] Otsu, N. 1979 Threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man, and Cybernetics* **9** 62-66.
[14] Sharma N, Ray AK, Sharma S, Shukla KK, Pradhan S, Aggarwal LM. 2009 Segmentation of medical images using simulated annealing based fuzzy C Means algorithm. *Int J Biomed Engg Technol.* **2** 260-278.
[15] Taylor, C. 1997 Classification and kernel density estimation. Vistas in Astronomy **41** 411-417.
[16] Vapnik, V. N. 1995 The nature of statistical learning theory.(New York, Springer-Verlag, Inc.).
[17] Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T., and Vapnik, V. 2000 Feature selection for svms. *Advances in Neural Information Processing Systems* (MIT Press) **13** 668-674.
[18] Zaidi H, Ruest T, Schoenahl F, Montandon ML 2006 Comparative assessment of statistical brain MR image segmentation algorithms and their impact on partial volume correction in PET. *Neuroimage*. **32** 1591-1607.
[19] Zhang, H., and Su, J. 2008 Naive bayes for optimal ranking. *J. Exp. Theor. Artif. Intell.* **20** 79-93.
[20] Boser, B. E., Guyon, I. , and Vapnik, V. 1992. A training algorithm for optimal margin classifiers. Proceedings of the Fifth Annual Workshop on Computational Learning Theory, 144-152.
[21] Cortes, C., and Vapnik, V. 1995. Support-vector network. Machine Learning, **20** 273-297

.