Voxel-based Discriminant Map Classification on Brain Ventrices for Alzheimer’s Disease

Jingnan Wang\textsuperscript{a,b} and Gerard de Haan\textsuperscript{a,b} and Devrim Unay\textsuperscript{a} and Octavian Soldea\textsuperscript{c} and Ahmet Ekin\textsuperscript{a}

\textsuperscript{a}Philips Research Europe, High Tech Campus 36, Eindhoven 5656 AE, The Netherlands; \textsuperscript{b}Eindhoven University of Technology, Den Dolech 2, Eindhoven 5600 MB, The Netherlands; \textsuperscript{c}Sabanci University, Orhanli-Tuzla, 34956, Istanbul, Turkey

\textbf{ABSTRACT}

One major hallmark of the Alzheimer’s disease (AD) is the loss of neurons in the brain. In many cases, medical experts use magnetic resonance imaging (MRI) to qualitatively measure the neuronal loss by the shrinkage or enlargement of the structures-of-interest. Brain ventricle is one of the popular choices. It is easily detectable in clinical MR images due to the high contrast of the cerebro-spinal fluid (CSF) with the rest of the parenchyma. Moreover, atrophy in any periventricular structure will directly lead to ventricle enlargement. For quantitative analysis, volume is the common choice. However, volume is a gross measure and it cannot capture the entire complexity of the anatomical shape. Since most existing shape descriptors are complex and difficult-to-reproduce, more straightforward and robust ways to extract ventricle shape features are preferred in the diagnosis. In this paper, a novel ventricle shape based classification method for Alzheimer’s disease is proposed. Training process has to be performed to generate two probability maps for two training classes: healthy controls (HC) and AD patients. By subtracting of the HC probability map from the AD probability map, we get a 3D ventricle discriminant map. Then a matching coefficient has been calculated between each training subject and the discriminant map. An adjustable cut-off point of the matching coefficients has been drawn for the two classes. Generally, the higher the cut-off point has been drawn, the higher specificity can be achieved. However, it will result in relatively lower sensitivity and vice versa. The benchmarked results against volume based classification show that the area under the ROC curves for our proposed method is as high as 0.86 compared with only 0.71 for volume based classification method.

\textbf{Keywords:} shape descriptors, brain ventricle, MRI, probability map, discriminant map, Alzheimer’s disease, classification, cross validation

\section{1. INTRODUCTION}

Dementia is a brain disorder that severely affects a person’s ability to carry out daily activities. The most common form of dementia is Alzheimer’s Disease (AD), accounting for about 74 percent of the cases in North America (where AD was the seventh leading cause of death in 2007 and the fifth leading cause of death for those over age 65\textsuperscript{1}), 61 percent in Europe, and 46 percent in Asia. Alzheimer’s disease initially involves the loss of short-term memory and is followed by a progressive cognitive decline. In its final and most severe stage, the patient is unable to perform any simple task and needs constant supervision. On average, the disease leads to death in about 7-10 years, and up to 20 years after the onset of the symptoms.\textsuperscript{2}

Today the only definite way to diagnose AD is to find out whether there are plaques and tangles in brain tissue.\textsuperscript{1} To investigate the molecular content of the brain tissue, however, doctors must wait until they do an autopsy. Therefore, doctors can only make a diagnosis of “possible” or “probable” AD while the person is still alive. Currently, the primary tests for AD diagnosis are cognitive tests. These tests measure various cognitive functions, such as memory, attention, orientation, language, and learning. The main appeal of these tests, such as mini-mental state examination (MMSE),\textsuperscript{3} is their ease of administration, but the results can be subjective and be affected by patient’s mental and physical state at the time of the test. In the diagnosis of AD, although cognitive test scores are powerful, structural imaging (MR, CT) is useful and is required for various clinical purposes:\textsuperscript{4} 1) to eliminate other possible causes, such as tumor or blood clots, for the low cognitive score, 2) to confirm that the patient does not have other forms of dementia, such as vascular and fronto-temporal,
3) because some patients, especially of higher education and intelligence, are able to hide cognitive deficits in the tests for a long time, image-based analysis can detect AD earlier by quantifying the brain structural changes.

Many studies have shown a correlation between the progression of the disease and the decrease of volume in periventricular structures, such as the hippocampus and the amygdala.\(^5-7\) One of the consequences is the enlargement of the brain ventricles.\(^8-9\) The ventricles are filled with CSF, a watery solution that provides physical and nutritional support to the brain. It enlarges at the expense of atrophy resulting from neuronal loss. In Fig.1, ventricles are visible as central hyper-intense regions for a normal control and an AD patient.

In literature\(^10-13\) many types of shape descriptors have been described. Most existing shape representation techniques use complex and difficult-to-implement shape descriptors. However, in medical applications, experts prefer more easily and robustly extractable features in their diagnosis. Thus the goals of this work are: 1) to develop new shape representation methods that will allow us to easily detect and quantify the differences in ventricle shapes due to the disease, 2) to verify that they have better performance than volume in the classification of AD patients and normal controls.

## 2. MATERIALS

### 2.1 Subjects

Our data set consists of 15 patients (age 70-80) with moderate AD diagnoses and 32 age-matched healthy controls. All subjects are right-handed. The demographic characteristics of the patients and normal controls are summarized in Table 1.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Controls (N=32)</th>
<th>Moderate (N=15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean SD</td>
<td>Mean SD</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>73.38 2.32</td>
<td>75 2.88</td>
</tr>
<tr>
<td>MMSE</td>
<td>29 1</td>
<td>21.53 3.67</td>
</tr>
<tr>
<td>Male/Female</td>
<td>8/24</td>
<td>6/9</td>
</tr>
</tbody>
</table>

### 2.2 Image Acquisition and Preprocessing

All data are T2 MR scans and are acquired on a Philips Intera 1.5T whole body scanner at Leiden University Medical Center with parameters (TR/TE:3000/120 ms, FLIP: 90, 220 mm FOV, 3 mm slice thickness, no slice gap and 256*256 matrix). A 3D MR data example is given in Fig.2.

For a selected slice, Fig.3 shows the key steps of the image preprocessing. Registration of the MR data into a common coordinate system has been made. Followed by the extraction of the CSF region by using a clustering-based segmentation algorithm with \( k = 3 \) clusters (background, CSF, and combined white and gray matter).\(^14\) From the CSF map, we extract ventricle map with spatial (brain atlas) and morphological constraints.
3. METHODS

In clinics, the volume of the ventricles, is used qualitatively or quantitatively in the diagnosis of AD. While volume can provide some indication of normal variation and anomaly, it is fairly crude and does not capture the entire complexity of the anatomical shape. The large overlap between the two volume distributions makes the measurement unsuitable as a biomarker of the disease. As shown in Fig.4, the volume range of the healthy controls is very wide. Although overall the ventricle volume of the AD patients (which is 32592.8 on average) is larger than that of the healthy controls (which is 23937.9 on average), there is no clear cut-off point for the two groups. If utilized properly, medical images can provide highly detailed shape information for analysis of morphological variability within a single population, or among different groups of subjects. Our proposed discriminant map classification is based on this 3D ventricle shape variability in a comprehensive way. It can be divided into three major steps as shown in Fig.5.

- Computing Discriminant Map
Figure 4. Volume distribution of the data set. The average volume of the healthy control group is 23937.9 compared with 32592.8 of Alzheimer’s patient group.

Figure 5. Schematic Diagram

Each binary ventricle slice can be interpolated into a matrix of 1s and 0s, in which 1 stands for the ventricle pixel in the binary image and 0 stands for the non-ventricle pixel. Before discriminant map, two probability map for each training class have to be calculated. As shown in Fig.6, the training set is first divided into two groups: Alzheimer’s patient group and healthy control group. An averaged brain atlas, or we called probability map, for each group is derived from the training data set, in which the value of the each pixel stands for the frequency of the ventricle pixels in this particular position. Fig.7 gives several examples of generated probability maps. Although the probability map of the HC group is similar to that of the AD group, there are more enlarged ventricles and greater atrophy of cortical gray matter in the patients with Alzheimer’s disease. By subtracting of the HC probability map from the AD probability map, we get the discriminant map (Fig.8).

- Setting Classifier Parameter
When the discriminant map $D(x,y,z)$ is available, matching coefficient ($C$) between each training input $S^i_{HC/AD}(x,y,z)$ and the discriminant map has to be calculated as a similarity measure:

$$C^i_{HC/AD} = \sum_{x,y,z} D(x,y,z) \cdot S^i_{HC/AD}(x,y,z)$$

The calculated matching coefficients (Fig.9) from the training set are used as a reference to get the optimized cut-off point for two groups. The optimized cut-off point can be defined differently to satisfy different purposes. Generally, higher cut-off point will lead to higher specificity, while lower cut-off point gives higher sensitivity. If sensitivity and
specificity are treated with equal weights, trade-offs have to be made in the training process to find out the cut-off point that brings about minimum combined errors for both groups.

Figure 8. Discriminant map generated from the training set

Figure 9. Illustration of the calculated matching coefficients distribution for training set

4. RESULTS

To measure the performance of a medical test, the concept sensitivity and specificity are often used. When some people have the disease, and if the test says they do. They are called true positives (TP). Some have the disease, but the test claims they don’t. They are called false negatives (FN). Some don’t have the disease, and the test says they don’t - true negatives (TN). Finally, we might have healthy people who have a positive test result false positives (FP). Thus, the number of true positives, false negatives, true negatives, and false positives add up to 100% of the set.

- Sensitivity (TPR) is the proportion of people that tested positive of all the positive tested; that is \( \frac{TP}{TP + FN} \).
It can be seen as the probability that the test is positive given that the patient is sick. The higher the sensitivity, the fewer real cases of diseases go undetected.

- **Specificity (TNR)** is the proportion of people that tested negative of all the negative people tested; that is \( \frac{TN}{TN + FP} \). As with sensitivity, it can be looked as the probability that the test is negative given that the patient is not sick. The higher the specificity, the fewer healthy people are labeled as sick.

The relationship between sensitivity and specificity can be visualized using the Receiver Operating Characteristic (ROC) curve. It is a graphical plot of the *Sensitivity* versus \( 1 - \text{Specificity} \) as its discrimination threshold varied. The area under ROC curve (AUC) \(^{15}\) is often used as a summary statistic for the ROC curve. The interpretation of the AUC is that: the higher the AUC, the better, with 0.50 indicating random performance and 1.00 denoting perfect performance.

![ROC curve](image.png)

**Figure 10.** ROC curves for two classification methods, with AUC of 0.86 for voxel-based probability map classification method and AUC of 0.71 for volume based classification

Due to our relatively small data size, leave one out cross validation is applied here: one case is left out of the training set and then used as a test set. Repeated for all the cases in the data set, this yields an estimate of the generalization accuracy of the method. The classification results of the proposed discriminant map based classification benchmarked against volume is given in Fig.10. The ROC curve of the probability map is dominant the curve of volume. With the given sensitivity, the specificity by the discriminant map based classification is always higher than that of volume. And with the given specificity, probability map based classification always achieves higher sensitivity. The AUC of the proposed voxel-based probability map classification is as high as 0.86, while only 0.71 for that of the volume based classification method. There are always trade-offs to achieve high sensitivity or high specificity. Thus in our training process, the threshold can be set for different goals.

**5. DISCUSSION AND CONCLUSIONS**

In this paper, we proposed a novel voxel-based discriminant map classification method on brain ventricles for Alzheimer's disease. It is an easy-to-compute alternative to the commonly-used volume measure. Unlike volume and other conventional shape descriptors, the discriminant map is derived from the complete 3D ventricle segmentation maps. No information of any single pixel is missing. Training process has to be performed to generate two probability maps for two training classes: healthy controls and AD patients. By subtraction of the HC probability map from the AD probability map, we get a discriminant map. Then a matching coefficient has been calculated between each training subject and the discriminant map. An adjustable cut-off point for the two classes has been drawn as the classification criteria for the test input. Generally,
the higher the cut-off point has been drawn, the higher specificity can be achieved. However, it will result in relatively lower sensitivity and vice versa. The benchmarked results against volume based classification show that the area under the ROC curves for our proposed method is as high as 0.86 compared with only 0.71 for volume based classification method.

Based on the test results, we conclude that the proposed voxel-based discriminant map classification method can be an easily-implemented alternative to volume for its good classification ability. As a further work, we are planning to confirm the results in a larger data set. Besides, anterior and posterior parts of the ventricles can also be taken into consideration separately to generate regional discriminant maps for further classification.

REFERENCES


