

TEXTON-BASED DIAGNOSIS OF ALZHEIMER'S DISEASE

*Pedro Morgado**, *Margarida Silveira**, *Durval Campos Costa^{†‡}* and
the Alzheimer's Disease Neuroimaging Initiative

*Instituto Superior Técnico, Institute for Systems and Robotics, Lisbon, Portugal

[†] Nuclear Medicine, Champalimaud Clinical Center, Champalimaud Foundation, Lisbon, Portugal

[‡]HPP Medicina Molecular, SA, Porto, Portugal

ABSTRACT

The textural content of FDG-PET brain images has been shown to be useful for the diagnosis of Alzheimer's disease (AD) and Mild Cognitive Impairment (MCI). In this paper, we investigate the use of the textons method [1], a powerful texture extraction procedure that uses a full statistical representation of the response of the image to a set of filters. We also extend the MR8 filter bank used in [1] to 3D in order to match the dimensionality of FDG-PET images, while maintaining important properties such as invariance to rotation and a low dimensionality of the filter response space. We propose two methods to tackle difficulties inherent to the extraction and classification of texture from images whose appearance varies over space and to the fact that most regions of the image are not affected by AD or MCI. The first method selects only the voxels with the most discriminative filter responses, while the second method focuses on brain regions manually labeled by an expert physician. Experiments showed that the proposed approaches outperformed the more common one that uses voxel intensities directly as features both in the diagnosis of AD and MCI. It was also observed that the discriminative power of certain brain regions increased significantly when the texton based analysis was performed.

Index Terms— Alzheimer's disease, Mild cognitive impairment, Texture analysis, Textons, Classification

1. INTRODUCTION

Alzheimer's disease is a neurological disorder for which no cure is currently available, causing long-term memory loss and affecting other cognitive abilities such as reasoning or planning [2]. As the disease progresses, these symptoms are aggravated and bodily functions are lost, leading eventually to death. Therefore, early detection of AD, when the brain is still not too deeply damaged, is essential to improve quality of life and extend life expectancy. However, the diagnosis of the condition that is typically associated with this early stage, Mild Cognitive Impairment (MCI) [3], is still troublesome because, among other difficulties, symptoms at this stage are often confused with "age related" issues [4].

This work was supported by Fundação para a Ciência e Tecnologia (FCT/MCTES) through the ADIAR project (PTDC/SAU-ENB/114606/2009). Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report.

Neuroimaging is one of the tools that is used to improve diagnostic accuracy. A particular technique that holds relevant information about both AD and MCI is the Positron Emission Tomography (PET) using fluorodeoxyglucose (FDG) as the tracer. This type of image measures at each location the consumption of glucose, and thus it is linked with brain activity and consequently with AD [5].

Consequently, many diagnostic systems based on FDG-PET images have been proposed, with most of them focusing on voxel intensity (VI) features, either extracted from the whole brain [6] or from regions of interest [7]. However, texture features have also been successfully used in the past. For instance, in [8], histograms of oriented gradients were used to extract textural information which were then fed to a Support Vector Machine (SVM) for the final diagnosis. In [9], a similar approach was conducted, but a 3D generalization of Local Binary Patterns was used as the texture descriptor. In this work, we also use the textural information of FDG-PET images, but we perform a texton-based texture analysis.

The textons method [1, 10] is a powerful texture extraction procedure that has raised much interest recently, owing its success mainly to a full statistical representation of the responses to a predefined set of filters. Moreover, textons have already been successfully applied to a variety of other problems, from very specific tasks, such as the detection and grading of lymphocytic infiltration in breast cancer histopathology [11] or face recognition [12], to more general problems such as image segmentation [13] and image retrieval [14], but to our knowledge it had never been applied to the analysis of PET images for the diagnosis of AD.

We based our diagnostic system on the method proposed by Varma and Zisserman (VZ) [1]. However, the VZ texton method was proposed for the analysis of full images of different textures and in the case of our FDG-PET analysis only a few regions are relevant for classification. Thus, joining the filter responses from the whole image may reduce the diagnostic accuracy. Therefore, we propose two methods based on texton analysis that circumvents this issue. One of them focuses only on the statistically most discriminative voxels, while the other one performs multiple independent texton analysis on different regions of the brain. We also compare the two proposed systems with a third, more common approach, that uses the voxel intensities of FDG-PET images directly as features.

The remaining of this paper is organized as follows. First, we present in section 2 the texton classification method as it was originally proposed in [1], as well as the 3D extension of the filter bank that we used in this work. Then, in section 3, we propose two approaches to deal with the problem at hand. The experimental setup and the results are described and discussed in section 4, and the most important conclusions are summarized in section 5.

2. TEXTURE CLASSIFICATION USING TEXTONS

Texture classification has been an active research field, with several methods being proposed to describe and eventually to classify texture images. In this work, the extraction of textural information was based on the algorithm proposed by Varma and Zisserman [1], which is described in the next section. Due to the 3D nature of FDG-PET images, the bank of filters used by this method had to be modified as explained in section 2.2.

2.1. Basic approach

The Varma and Zisserman (VZ) approach is divided into two stages: learning and classification.

During the learning stage, a fixed number of images of each class, n_I , are randomly selected and convolved with a filter bank, generating a set of filter responses. Then, the filter responses at all pixel positions in those images are clustered into n_T clusters, or textons, using the standard *k-means* algorithm. This clustering procedure is repeated independently for each one of the n_C classes giving rise, at the end, to a dictionary of $n_T \times n_C$ textons. The learning stage ends by generating for each training image a model that will represent it from that point forward. Such models are built by labeling each position of the image with the closest texton, and then computing the normalized histogram of those labels. Each model is therefore a vector with $n_T \times n_C$ elements where the n -th element represents the probability of the filter responses at any position of the image being most reliably represented by the n -th texton of the dictionary.

During the classification stage, the response of the novel image to the filter bank is computed, and a model is constructed in the same way as with training images. The classification is then carried out with a nearest neighbor classifier, i.e., by assigning to the new image the class of the nearest training model. A χ^2 distance was used in the VZ classifier to measure the similarity between histograms.

In addition, following the original approach, three preprocessing steps should be carried out before the learning stage. The intensity of the input images should be normalized to have zero mean and unit variance, each filter should also be normalized so that it has unit l_1 norm, and the filter response at each voxel position \mathbf{x} should also be normalized according to the following equation (motivated by Webber's law):

$$\mathbf{F}(\mathbf{x}) \leftarrow \frac{\mathbf{F}(\mathbf{x})}{\|\mathbf{F}(\mathbf{x})\|_2} \log \left(1 + \frac{\|\mathbf{F}(\mathbf{x})\|_2}{0.03} \right). \quad (1)$$

For more information on these preprocessing procedures, the interested reader should refer back to [1] and references therein.

2.2. Three-dimensional filter bank

Originally, the VZ classifier was proposed with a set of 2D filters known as the Maximum Response 8 (MR8) filter bank. The MR8 filter bank consists of a Gaussian, a Laplacian of Gaussian, and edge and line filters, each at 3 pairs of scales and 6 orientations. Although this results in 38 filter responses, only the maximum across all 6 orientations is kept for the edge and line filters and therefore only 8 filter responses are used. The goal of this maximization step is to achieve rotation invariance, which is important when images are allowed to appear with any orientation and to reduce the dimensionality of the filter response space where textons are searched for. In the problem at hand, the latter advantage is more significant than the former since all PET images were previously aligned.

In this work, since PET images are three dimensional, a 3D version of the MR8 filter bank was used. More concretely, the filter bank was composed by a 3D Gaussian filter and its Laplacian:

$$G(\mathbf{x}, \sigma) = K \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^3 x_i^2 \right\}, \quad (2)$$

$$\nabla^2 G(\cdot) = -\frac{K}{\sigma^2} \left(1 - \frac{\sum_{i=1}^3 x_i^2}{\sigma^2} \right) \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^3 x_i^2 \right\}, \quad (3)$$

where K is a normalization factor and $\mathbf{x} = (x_1, x_2, x_3)$ is the vector of spatial coordinates, and three other types of filters at 3 triplets of scales $(\sigma_1, \sigma_2, \sigma_3)$ and 61 orientations. Namely, 3D edge filters, plane filters and line filters where an exemplar of each type of filter is given by:

$$\frac{\partial G(\cdot)}{\partial x_3} = -\frac{K \cdot x_3}{\sigma_3^2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^3 \frac{x_i^2}{\sigma_i^2} \right\}, \quad (4)$$

$$\frac{\partial^2 G(\cdot)}{\partial x_3^2} = -\frac{K}{\sigma_3^2} \left(1 - \frac{x_3^2}{\sigma_3^2} \right) \exp \left\{ -\frac{1}{2} \sum_{i=1}^3 \frac{x_i^2}{\sigma_i^2} \right\}, \quad (5)$$

$$\sum_{i=1}^2 \frac{\partial^2 G(\cdot)}{\partial x_i^2} = -K \sum_{i=1}^2 \left(\frac{\sigma_i^2 - x_i^2}{\sigma_i^4} \right) \exp \left\{ -\frac{1}{2} \sum_{i=1}^3 \frac{x_i^2}{\sigma_i^2} \right\}, \quad (6)$$

respectively. Despite the large number of filters, since only the maximum across different orientations is kept, the final filter response space has only 11 dimensions $(2 + 3 \times 3)$ which is still small enough for a reliable clustering. One exemplar of each type of filter is shown in Fig. 1.

3. ALZHEIMER'S DISEASE DIAGNOSIS USING TEXTONS

In the previous section, the basic texture classification algorithm was presented. As mentioned before, the original VZ classifier was designed to recognize a set of pre-learned texture classes from texture images, i.e. from images where any patch of sufficiently large dimension has similar appearance and no distracting background.

The diagnosis of AD from FDG-PET images presents different properties that needed to be dealt with in order to achieve good generalization. First, only a few regions within the original volume present distinct appearances across individuals of different classes. Second, the appearance of the image is not uniform over the entire volume. Also, the 3D nature of PET images and the consequent larger number of voxels makes the computational needs a constraint. On the other hand, all images share the same orientation since they were registered to the same anatomical space, their intensities are normalized and there are no problems regarding viewing angles.

In order to take these properties into account, we propose the following two approaches for the diagnosis of AD.

3.1. Texton analysis of the most discriminative voxels

Most regions of the FDG-PET images have similar appearance across all subjects despite the inter subject brain variability. Therefore, if we computed the final histograms using the labels of the textons associated with the entire image, the contribution to the final model of the smaller regions with discriminative appearance would be mixed together with the labels of the remaining non-discriminative regions and the generalization ability would be reduced. Consequently, we should expect improved results if we

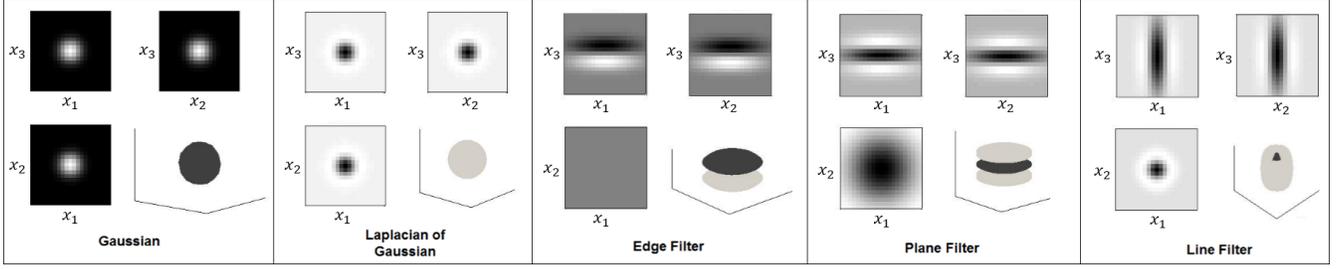


Fig. 1. Filter bank. An example of each type of filter is given. For each filter, three perpendicular cuts are shown: the top-left, top-right and bottom-left images represent the cuts through the planes $x_2 = 0$, $x_1 = 0$ and $x_3 = 0$, respectively. In the bottom-right images, isosurfaces of the kernels are presented.

restrict the texture analysis only to the most relevant regions of the brain.

Several measures of relevance can be found in the literature, but since we are searching for the best voxels, which are now represented by 11-dimensional vectors, measures such as mutual information that require the full estimation of the underlying distribution should fail due to the sparsity of training data. In this work, we opted for a wrapper method instead, i.e. we used a simple classifier to score each voxel independently. More specifically, we used a 9-nearest neighbor classifier and scored each voxel based on its accuracy computed through a cross-validation procedure. After scoring all voxels, we selected the best n_V and used their filter responses to build the dictionary and extract the models. Then, in addition to the nearest neighbor classifier based on a χ^2 distance, a few others were tested in order to search for the best generalization, as explained in section 3.3.

Just for the sake of comparison, we also used the same number of VIs directly for classification (instead of the histograms of textons), but in this case the n_V voxels were selected based on the mutual information between its intensity and the class label.

Note that these discriminative selection schemes can only be applied because the images were previously registered to a common space (the Talairach space to be precise) and, thus, equivalent voxels correspond to the same anatomical position.

3.2. Independent texton analysis of ROIs

Different regions of the PET volume may exhibit distinct appearances but only some of them are associated with AD related changes. In order to deal with this characteristic, we analyzed 7 disjoint regions independently. These regions were manually labeled by an expert physician on the average PET volume of Normal Controls (NC), and only regions associated with AD related changes were analyzed, namely: 1) left and right lateral temporal; 2) left and right mesial temporal; 3) inferior frontal gyrus; 4) inferior anterior cingulate; 5) left and right dorsolateral parietal; 6) superior anterior cingulate; 7) posterior cingulate and precuneus.

Image classification using this approach was performed in the following way. First, the texton dictionary was built, and the models computed using only the filter responses at locations belonging to a single region. This procedure was repeated for each region, resulting in 7 different dictionaries and 7 models for each one of the training and testing images. Then, two classification schemes were compared. In the first, all histograms were concatenated and several classifiers tested (see section 3.3), and, in the second, each region was classified independently and the predictions were combined using the majority voting rule.

Once again, we also compared this approach with one that uses the intensity of the voxels inside each region directly as features.

3.3. Classifiers

Before describing the different classifiers that were used, let us introduce the notation. Each subject is represented by an N -dimensional feature vector $\mathbf{f}^{(p)} = (f_1^{(p)}, \dots, f_N^{(p)})$ and its class label $y^{(p)}$, where the superscript p indexes subjects in the database. This is valid whether features represent bins of the histogram or voxel intensities.

As for the classifiers, two different types were tested: k -Nearest Neighbors and Support Vector Machines.

3.3.1. k -Nearest Neighbors

A nearest neighbor classifier was originally proposed for the classification of the models extracted from a texton analysis [1]. Therefore, we also tested this approach, but we used the more general k -nearest neighbors (k -NN) classifier with the parameter k being selected by cross-validation.

The χ^2 distance (equation (7)) was preferred over the standard l_2 distance because it is more suitable to measure similarities between histograms:

$$d_{\chi^2}(\mathbf{f}^{(p)}, \mathbf{f}^{(q)}) = 2 \sum_{i=1}^N \frac{(f_i^{(p)} - f_i^{(q)})^2}{f_i^{(p)} + f_i^{(q)}}. \quad (7)$$

3.3.2. Support Vector Machine

A Support Vector Machine [15] is an algorithm that searches for the class separation hyperplane with maximum margin, whether it is on the original or some other (typically higher dimensional) space. An SVM is also able to deal with non separable data by using the soft margin concept, which allows for mislabeled examples to occur while penalizing them. On the other hand, the use of kernel functions enables us to map (implicitly) the input patterns into a higher dimensional space, and consequently to learn non-linear separation surfaces in the original space.

In this work, in addition to the linear kernel:

$$K_L(\mathbf{f}^{(p)}, \mathbf{f}^{(q)}) = \mathbf{f}^{(p)} \cdot \mathbf{f}^{(q)}, \quad (8)$$

we also tested the generalized histogram intersection (GHI) kernel [16] because it is able to measure the degree of similarity between two histograms. The GHI kernel is defined as:

$$K_{GHI}(\mathbf{f}^{(p)}, \mathbf{f}^{(q)}) = \sum_{i=1}^N \min\left(|f_i^{(p)}|^\beta, |f_i^{(q)}|^\beta\right), \quad (9)$$

where the parameter β was introduced to add flexibility to the underlying mapping function and can be used to improve generalization.

4. EXPERIMENTAL RESULTS

We now describe the image database that was used herein and the experiments that were conducted to validate the two approaches proposed in this paper.

4.1. ADNI Database

All PET images were retrieved from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database, but the following class dependent restrictions were imposed to the Clinical Dementia Rating (CDR) score of each subject: 0 for healthy controls, 0.5 for MCI patients and 0.5 or higher for AD patients, resulting in a dataset composed by 59, 104 and 70 subjects, respectively. Table 1 summarizes important clinical and demographic information about each group.

The retrieved data had already undergone a series of preprocessing steps in order to minimize differences between images and thus allowing voxel-wise comparisons. More specifically, every PET image was reoriented (such that the anterior-posterior axis of the subject is parallel to the AC-PC line), normalized in its intensity, and smoothed to a uniform standardized resolution [17]. Finally, they were all warped to the Tailarach brain atlas [18].

Table 1. Characteristics of each group. Format: Mean (Standard Deviation). MMSE stands for Mini Mental State Exam.

Attributes	AD	MCI	NC
N ^o of subjects	59	104	70
Age	78.0 (6.6)	77.1 (7.2)	77.7 (4.9)
Sex (% of Males)	55.9	64.4	65.7
MMSE	19.6 (5.1)	26.3 (3.1)	29.2 (0.9)

4.2. Experimental setup

The proposed approaches were evaluated on two diagnostic problems, AD vs. NC and MCI vs. NC, under several conditions.

The filter bank that was used was a 3D version of the MR8 filter composed by 5 types of filters at several scales and orientations, as previously described in section 2.2. A single scale ($\sigma = 2$) was used for the Gaussian filter and its Laplacian (no rotations were needed since they are already rotation invariant). The 3D edge and plane filters were included at 3 triplets of scales, $(\sigma_x, \sigma_y, \sigma_z) = \{(1.5, 1.5, 0.5), (3, 3, 1), (6, 6, 2)\}$, as well as the line filter, $(\sigma_x, \sigma_y, \sigma_z) = \{(0.5, 0.5, 1.5), (1, 1, 3), (2, 2, 6)\}$. The last 3 types of filters were replicated with multiple orientations. In fact, equations (4) to (6) only allow us to build the version of the filters with rotational symmetry around the z -axis. The remaining versions were obtained by rotating this one φ radians around the z -axis, first, and then θ radians around the y -axis. In addition to the initial orientation, we also added to the filter bank filters rotated with all combinations of $\theta = \{\frac{\pi}{6}, \frac{2\pi}{6}, \dots, \frac{5\pi}{6}\}$ and $\varphi = \{0, \frac{\pi}{6}, \dots, \frac{11\pi}{6}\}$.

The number of images, n_I , that were initially used for clustering was fixed to 50. We performed a few experiments to study this parameter but, as expected, it did not influence much the performance of the recognition system, unless a very small number of images were used. These experiments are not presented here due to space concerns. The number of textons per class, n_T , and the number of selected voxels, n_V , were also studied as explained in the next section.

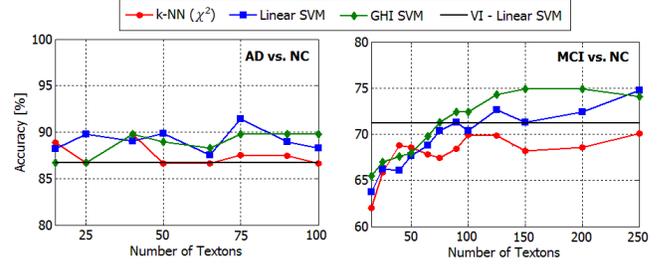


Fig. 2. Influence of the number of textons per class on the diagnostic accuracy for the problems: AD vs. CN (left) and MCI vs. CN (right).

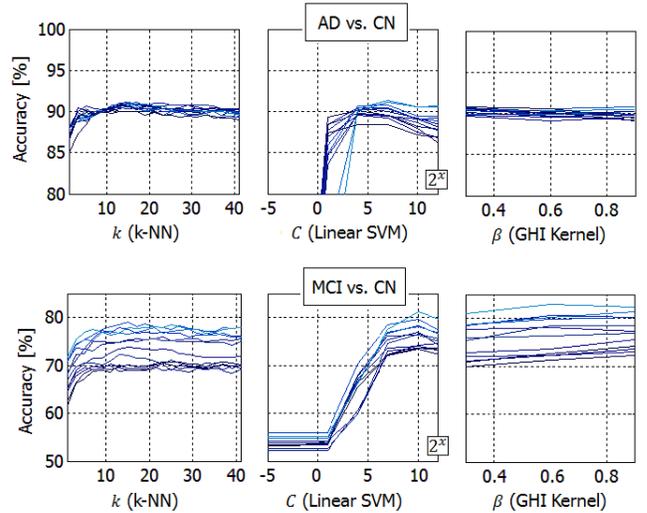


Fig. 3. Influence of classifier parameters on the diagnostic accuracy attained on the validation sets during the nested cross-validation for the problems: AD vs. NC (top row) and MCI vs. NC (bottom row). The color encodes the number of textons, with darker tones meaning smaller numbers.

Regarding the parameters associated with the classifiers, they were all tuned inside a nested cross-validation procedure. The allowed values were the following: number of neighbors of the k -NN algorithm, $k \in \{1, 3, \dots, 41\}$; parameter C of the SVM algorithm, $C \in \{2^{-5}, 2^{-3}, \dots, 2^{11}\}$; parameter β of the GHI kernel, $\beta \in \{0.2, 0.4, \dots, 1\}$. In addition to the classifiers presented in this paper, we also tested the Naive Bayes and the l_2 based k -NN classifiers but their performances were not as competitive.

As for the VI approach, we only present the results attained by a linear SVM. However, we also tested the Radial Basis Function (RBF) kernel, but worse performances were achieved.

4.3. Texton analysis of the most discriminative voxels

The first approach, where only the most discriminative voxels were used to build the texton dictionary and to compute the models, was studied extensively.

First, we studied the effect of the parameters of the system on the generalization ability on the two binary problems. We evaluated all combinations between a set of numbers of textons per class, $n_T \in \{15, 25, 40, 50, 65, 75, 90, 100\}$, and a set of numbers of selected voxels, $n_V \in \{1000, 2000, \dots, 9000\}$. Figure 2 shows the maximum diagnostic accuracy over the allowed numbers of selected

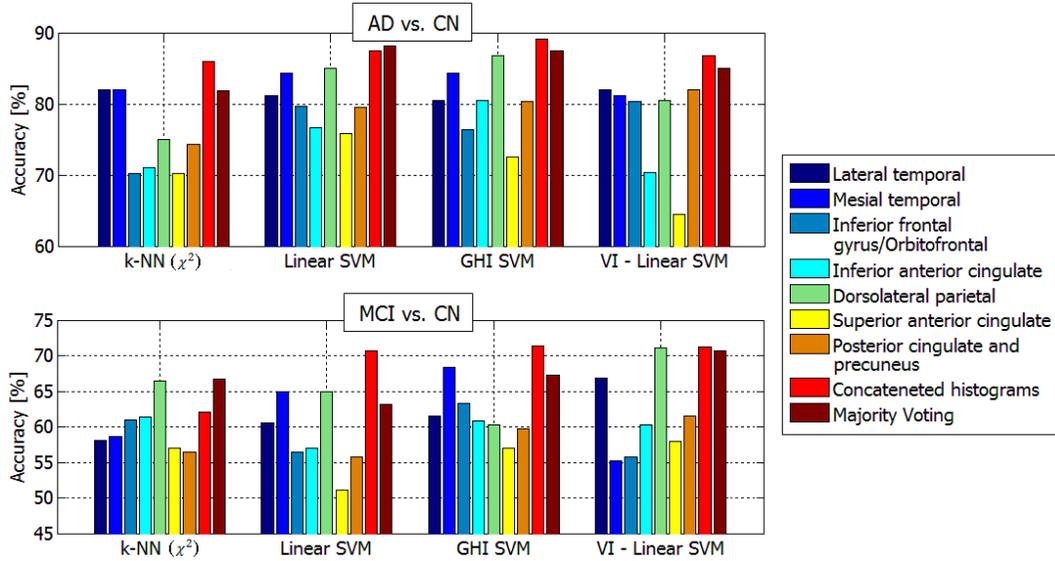


Fig. 4. Diagnostic accuracies achieved with an independent texton-based analysis of 7 brain regions for the problems AD vs. NC (on the top) and MCI vs. NC (on the bottom). Performances attained using each region individually, concatenated histograms and majority voting are shown. Also, the last group of results, “VI – Linear SVM”, shows, for comparison purposes, the results attained using VIs and the linear SVM for classification.

voxels as a function of the number of textons for both diagnostic problems (AD vs. CN and MCI vs. CN).

As can be seen, all texton-based classification schemes achieved at some point better performances than VI (black straight line) when discriminating between AD and NC, with the best score of 91.4% being reported by the linear SVM using dictionaries with 75 textons. As for the diagnosis of MCI, the texton analysis was largely aided by the increase of the dictionary size. Consequently, we subsequently allowed larger numbers of textons and achieved the best overall accuracy of 74.9% using 200 textons per class and an SVM with the GHI kernel for classification. These performances are significantly better than the 86.7% and 71.2% accuracies achieved by the VI approach on the AD vs. CN and MCI vs. CN problems, respectively.

Qualitatively, the impact of the number of textons on the generalization is consistent with the difficulty of the problem. In fact, the diagnosis of MCI is admittedly harder than the diagnosis of AD, which means that the underlying distributions of the filter responses among competitive classes are more similar, and thus refining the Voronoi partitioning of the filter response space, introducing more centroids, helps the system to find the smaller discriminative regions of the filter response space. On the other hand, the number of selected voxels did not affect considerably the generalization of the system as long as it was set high enough.

In what regards the type of classifier used for the diagnosis, our experiments led us to the conclusion that an SVM based classification generally leads to a significant improvement over the k -NN classifier initially proposed in [1], and that the GHI kernel only confirmed its theoretical advantage over the linear kernel on the task MCI vs. CN (see Fig. 2). In addition, although the parameters associated with classifiers were always tuned inside the nested cross-validation procedure, we also wanted to observe their influence on the final results. Fig. 3 gives us the general picture, where the top row is concerned with the diagnosis of AD and the bottom row with the diagnosis of MCI. Note that, each line, which was estimated as

the average accuracy attained on the validation sets, corresponds to a pair of parameter values (n_T , n_V). Two important observations can be drawn from the analysis of this figure. On one hand, we can see that, in these two specific problems, better accuracies might be achieved if we base our classification in the consensus decision of more than one neighbor (when a k -NN approach is used), in contrast with what was originally proposed in [1]. On the other hand, it can be observed that there is significant overfitting in the tuning of the classifier parameters on the MCI vs. CN problem, since large differences are visible between the best performances observed on the validation and test sets.

4.4. Independent texton analysis of ROIs

The second approach extracts textural information in seven regions of the brain, independently, using the texton analysis and, then, two classification schemes are tested. The first one concatenates the seven models (histograms of texton labels) of every subject and classifies the resulting feature vectors using one of the classifiers described in section 3.3. The second one performs the diagnosis directly from the models of each region and then combines the resulting seven predictions into a single, hopefully more reliable one through the majority voting rule. The most important results are shown in Fig. 4, where all experiments were conducted using a fixed number of textons per class $n_T = 90$. Note that, in addition to the final classification accuracies, we also present the performances that would be achieved if each region was used as the only source of information. In order to keep the comparison fair, we also evaluated the VI approach (with the linear SVM) using only voxels from each one of the selected regions independently. Several details should be emphasized.

First, regions such as dorsolateral parietal, lateral temporals and mesial temporals are already very discriminative by themselves, especially on the problem opposing AD patients to healthy controls. Nevertheless, the combination of regions achieved enhanced results

in almost all scenarios, with the best performances being achieved in the two tasks (89.1% and 71.4%, respectively) by concatenating region models and feeding them to an SVM with the GHI kernel. Also, it is worth noting that several significant improvements were achieved by the proposed method on isolated regions when compared to the equivalent VI experiment, for instance, on the dorsolateral parietal in the AD vs. CN problem or the mesial temporal in the MCI vs. CN, among others.

On the other hand, the SVM based classification proved, once again, to attain performances superior to the k -NN algorithm, in conformity with the results presented in the previous section, but not superior to the VI approach when diagnosing MCI. In fact, the differences between these two schemes are so small that they should be dismissed.

Finally, we can also notice that the approach discussed in the previous section (with a discriminative selection of image voxels) has better generalization ability than this one (with a selection based on ROIs) for both diagnostic problems. As mentioned before, the appearance of a PET image is not uniform over the entire volume (as in texture images) and, thus, less discriminative models may be constructed if the spatial information is completely disregarded. However, the partitioning used herein might not be optimal for this specific task, even though the delineated regions have been carefully selected. Therefore, automatic segmentation approaches, which are able to search for better partitions of the PET image, can be used in this context to possibly enhance the system's performance.

5. CONCLUSION

In this paper, we proposed the use of 3D textons for the diagnosis of AD and MCI. We presented two approaches to circumvent some difficulties associated with the use of a texture classification algorithm to perform the diagnosis from PET images. The first approach focused only on the most discriminative voxels, while the second focused on a set of manually labeled ROIs.

Experiments allowed us to conclude that both approaches outperformed the most common one, based on VIs, in the diagnosis of AD. As for the diagnosis of MCI, although both methods have reported higher accuracies, the differences were only significant when a discriminative selection scheme was used. Overall, we consider that the results attained in this work are convincing, especially when taking into account the small portion of the PET image that was used to conduct the texture analysis.

6. REFERENCES

- [1] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *International Journal of Computer Vision*, vol. 62, no. 1–2, pp. 61–81, 2005.
- [2] Z.S. Khachaturian and T.S. Radebaugh, *Alzheimer's Disease: Cause(s), Diagnosis, Treatment, and Care*, Taylor & Francis Group, 1996.
- [3] J.C. Morris, M. Storandt, and J. Miller, "Mild cognitive impairment represents early-stage Alzheimer disease," *Archives of Neurology*, vol. 58, no. 3, pp. 397–405, 2001.
- [4] G. Waldemar, B. Dubois, M. Emre, J. Georges, I. G. McKeith, M. Rossor, P. Scheltens, P. Tariska, and B. Winblad, "Recommendations for the diagnosis and management of Alzheimer's disease and other disorders associated with dementia: EFNS guideline," *European Journal of Neurology*, vol. 14, no. 1, pp. e1–e26, 2007.
- [5] K. Herholz, S. Carter, and M. Jones, "Positron emission tomography imaging in dementia," *The British Journal of Radiology*, vol. 80, pp. 160–167, 2007.
- [6] R. Chaves, J. Ramírez, J. M. Górriz, M. López, I. Álvarez, D. Salas-Gonzalez, F. Segovia, and P. Padilla, "SPECT image classification based on NMSE feature correlation weighting and SVM," in *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, 2009, pp. 2715–2719.
- [7] K. R. Gray, R. Wolz, S. Keihaninejad, R. A. Heckemann, P. Aljabar, A. Hammers, and D. Rueckert, "Regional analysis of FDG-PET for use in the classification of Alzheimer's disease," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, 2011, pp. 1082–1085.
- [8] E. Bicacro, M. Silveira, and J. S. Marques, "Alternative feature extraction methods in 3D brain image-based diagnosis of Alzheimer's disease," *Image Processing (ICIP'12), 2012 IEEE International Conference on*, pp. 134–137, 2012.
- [9] P. Morgado, M. Silveira, and J. Marques, "Diagnosis of Alzheimer's disease using 3D local binary patterns," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2013.
- [10] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *Int. J. Comput. Vision*, vol. 43, no. 1, pp. 29–44, June 2001.
- [11] A.N. Basavanahally, S. Ganesan, S. Agner, J.P. Monaco, M.D. Feldman, J.E. Tomaszewski, G. Bhanot, and A. Madabhushi, "Computerized image-based detection and grading of lymphocytic infiltration in her2+ breast cancer histopathology," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 3, pp. 642–653, 2010.
- [12] C. Zhong, Z. Sun, and T. Tan, "Robust 3D face recognition using learned visual codebook," in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, 2007, pp. 1–6.
- [13] J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and texture analysis for image segmentation," *Int. J. Comput. Vision*, vol. 43, no. 1, pp. 7–27, June 2001.
- [14] G. Liu and J. Yang, "Image retrieval based on the texton co-occurrence matrix," *Pattern Recognition*, vol. 41, no. 12, pp. 3521–3527, 2008.
- [15] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [16] S. Boughorbel, J. P. Tarel, and Nozha Boujemaa, "Generalized histogram intersection kernel for image recognition," in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, 2005, vol. 3, pp. III–161–4.
- [17] Alzheimer's Disease Neuroimaging Initiative, Laboratory of Neuro Imaging, UCLA, "PET pre-processing," [Online] Available: <http://adni.loni.ucla.edu/methods/pet-analysis/pre-processing/>, 2012.
- [18] Center for Alzheimer's Care Imaging and Research, "Summary of Neurostat processed images uploaded to LONI," [Online] Available: http://www.loni.ucla.edu/twiki/pub/ADNI/ADNIPostProc/UUtah_Analysis.pdf, 2007.