DEMENTIA DIAGNOSIS USING SIMILAR AND DISSIMILAR RETRIEVAL ITEMS

Devrim Unay[⋆] *Ahmet Ekin*[†]

[⋆]Electrical and Electronics Engineering, Bahcesehir University, Istanbul, Turkey
[†]Video Processing and Analysis Group, Philips Research, Eindhoven, The Netherlands
[⋆]devrim.unay@bahcesehir.edu.tr [†]ahmet.ekin@philips.com

ABSTRACT

Image-based disease diagnosis often requires radiologists' qualitative interpretations that are highly dependent on their levels of expertise, and physical and mental status. Automated image analysis tools can help radiologists increase the diagnosis accuracy by providing quantitative measures. Accordingly, this paper presents an automated method for dementia diagnosis using search and retrieval of brain MR (magnetic resonance) images. The main contributions of the method are 1)its generic decision-making based on similar and dissimilar cases retrieved from a database that can be used to diagnose various brain disorders, 2)it realizes dementia diagnosis utilizing a tailored version of histogram of oriented gradients (HOGs) as features, and 3)it achieves high performance in dementia diagnosis that is independent of the database size. Comprehensive experiments with real data showed that combining information from similar and dissimilar cases leads to improved diagnostic accuracy than using similar or dissimilar cases alone, accuracy diminishes with extreme quantization of the HOGs and with small databases, and our method achieves high performance with comparable sensitivity score to the most skilled experts at better specificity figures.

Index Terms— Image retrieval based dementia diagnosis, similar and dissimilar cases, Histogram of Oriented Gradients, brain MRI

1. INTRODUCTION

Dementia is the loss of cognitive abilities, such as decision making and memory, resulting mainly from neurodegeneration (deterioration of brain cells and their interconnections). An estimated 30 million people suffer from dementia, worldwide, and the number is expected to rise over 80 million by 2040 [1].

Neuroimaging, traditionally used to exclude treatable causes of dementia [2], has become essential for the examination of dementia patients over the past few years, due to the increasing need for more accurate diagnosis and prognosis. Therefore, practice guidelines recommend that neuroimaging should be performed for all dementia cases referred to the neurologist.

Image-based diagnosis of dementia, in the clinical settings, is often performed using qualitative measures, while the medical image analysis research has focussed on the automated computation of quantitative measures to improve the diagnosis and efficiency. In addition, as there exists high variation in the characteristics of different dementia types and their causes and progress are still not yet fully known, the comparison of multiple patients, their pathologies and progresses by search and retrieval systems should especially improve their diagnosis.

In the light of the above, this paper[1] presents an automated method for dementia diagnosis using search and retrieval of brain MR (magnetic resonance) images.

2. RELATED WORK

In differentiating dementia patients from cognitively normal subjects using image data, visual interpretation is the common practice in the clinical settings despite its large interrater variability [3]. To overcome subjectivity, researchers have focused on quantitative interpretations of images by employing region-of-interest (ROI) based approaches [4], and voxel-based morphometry (VBM) methods [5]. Due to the spatial limitation of ROI-based approaches and the sensitivity of VBM-based methods to systematic shape differences attributable to misregistration, recent focus has shifted to the use of machine learning methods [6][7].

Machine learning methods can provide more specialized and more accurate solutions after their parameters are optimized for a specific task. Unfortunately, any change in their task definition or the training dataset requires re-tuning of their parameters. Search and retrieval offers a promising alternative to overcome this drawback. The previous work, however, only addressed the generic problem of MR image retrieval [8][9]. In contrast, in this paper, we propose a search and retrieval method to support a diagnostic decision by ranking of the database entries and combining the information

[1]This work was supported in part by the Marie Curie Programme of EC under FP6 IRonDB project MTK-CT-2006-047217.

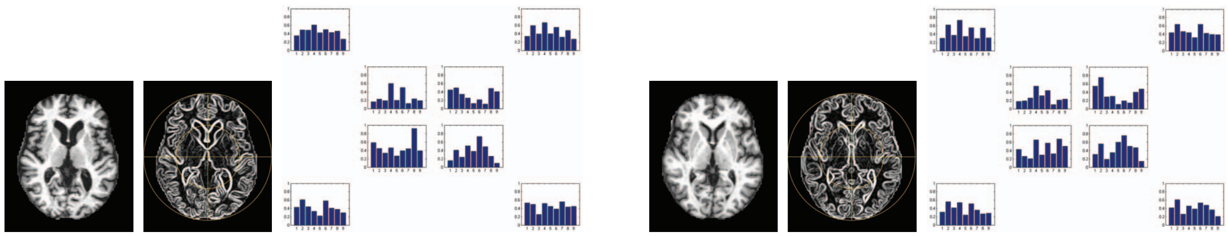

Fig. 1. HOGs computed for two subjects. Left-to-right: Original image, related gradient image partitioned into 4 angular and 2 radial regions, and the corresponding 9-bin histograms computed from each region are displayed for each subject.

coming from the similar and dissimilar returns.

3. DEMENTIA DIAGNOSIS FROM HOG-BASED BRAIN MR IMAGE RETRIEVAL

The search and retrieval method used for dementia diagnosis in this work consists of the pre-processing, feature extraction, similarity computation and decision steps. Pre-processing includes spatial alignment of the MR volumes and brain tissue extraction by Brain Extraction Tool [10]. The rest of the processing steps are explained in the following subsections.

3.1. Feature Extraction: Histogram of Oriented Gradients

Histogram of oriented gradients, or HOGs, are feature descriptors used in computer vision for object detection, such as human detection [11]. HOG represents objects by the occurrences of gradient orientation in local portions of an image. It computes its descriptors on a dense grid of uniformly spaced cells and benefits from overlapped local contrast normalization. These attributes provide HOGs with key advantages, such as invariancy to intensity variation, a problem in MR images. In contrast to the conventional HOGs image tesselation of a rectangular grid, we adapt HOGs to our application by using a polar grid fitted to the brain area. The polar grid is preferred over the rectangular version, because it conforms better to the brain shape. Figure 1 shows two brain MR images, the corresponding gradient images partitioned into 4 angular and 2 radial regions using the polar grid, and the related region histograms using 9 evenly spaced bins over 360°.

3.2. Similarity Computation and Decision

In our retrieval problem, query and target are represented by the histograms computed for each region, and the similarity is measured using the Bhattacharyya distance:

$$s_{q,t} = \sum_{\forall regions} \sum_{\forall bins} \sqrt{H_q H_t} \qquad (1)$$

Once similarity for each query-target combination is computed, they are sorted and used for the final diagnostic decision. While in the conventional systems decision is made

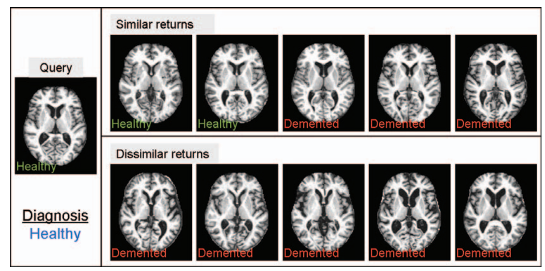

Fig. 2. A retrieval result showing similar and dissimilar items for a query, the true classes (ground truth), and the system's diagnostic decision.

based on the items similar to a query, we have observed that for our specific problem benefiting from dissimilar as well as similar items highly improved performance (Figure 2). A straightforward way of fusing the two information is to use a weighted combination scheme: Similar and dissimilar items contribute to the final decision with the weights of w and $1 - w$, respectively, given that $w \in [0, 1]$. Accordingly, the final class of the query, c_q, is computed as the weighted combination of the class values of the k_h highest and the k_l lowest retrieval returns. If the database consists of N items ranked in decreasing order of similarity as $i = 1, \ldots, N$ given the query, then

$$c_q = \left(w \frac{1}{k_h} \sum_{i=1}^{k_h} c_i \right) + \left((-1)(1-w) \frac{1}{k_l} \sum_{j=N-k_l+1}^{N} c_j \right) \qquad (2)$$

where $c = \{-1, +1\}$ is the class value of an item indicating its diagnosis.

4. EXPERIMENTS AND RESULTS

4.1. Subjects and Image Data

The database consists of 118 elderly patients (78 females) from OASIS [12]. In OASIS, patient information, MR acquisition, and tests to assess the cognitive performance of the subjects, such as the clinical dementia rating (CDR) scale [13], are recorded. CDR is an elaborate test to discern very mild dementia, where a score of 0 indicates no dementia, while higher scores show dementia with increasing severity. Based on the CDR scores, we grouped 77 subjects as cognitively normal controls, while 41 having dementia. Each group is matched on gender and age.

For each subject, a T1-weighted MR scan is collected using a 1.5T scanner with the following parameters: TR=9.7, TE=4.0, flip angle=10, voxel size 1mm$\times$ 1mm$\times$ 1.25mm, and matrix dimensions 256$\times$256.

4.2. Performance Evaluation

We measure the performance of a test by using the *F-measure*, which is the harmonic mean of *sensitivity* (probability of correctly diagnosing a condition) and *specificity* (probability of correctly identifying a non-diseased person). Note that higher F-measure values depict more accurate tests.

4.3. Results

The following results are achieved by evaluation in a leave-one-out fashion, where each subject in the database is used once as query with the rest forming the target set.

The MR scan of each subject in our database is composed of 30 informative slices for dementia (see [14] for slice matching across subjects). We analysed effects of various parameters. Figure 3 shows the effect of spatial partitioning on the retrieval performance. We observe that under and over tesselation result in lower performances. Partitioning into 4 angular and 2 radial regions generates the best retrieval performance, because this tesselation provides better separation of the cortical and deep-cortical structures of the brain.

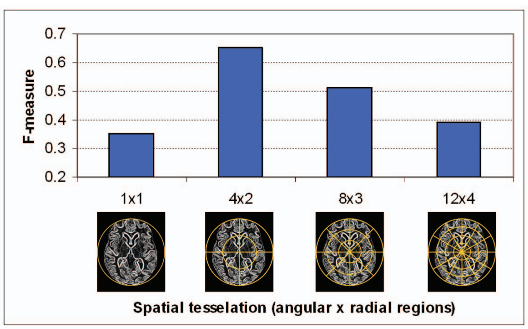

Fig. 3. Effect of spatial partitioning.

In addition, our analysis of the effect of HOG bin size on the retrieval performance showed that extreme values generally lead to under or over quantization of the histograms and thus lower performances. Quantization into 8 bins, similar to the best performing quantization observed in human detection [11], result in the best retrieval performance.

Figure 4 shows the performance when the number of similar and dissimilar items used for decision-making is varied. As observed, using 5 or 7 items from the highest and lowest ranked returns, which correspond to approximately 10% of the database, has lead to improved performance. Notice that these quantities correspond to the typical values for k used in the conventional k-nearest neighbour algorithm.

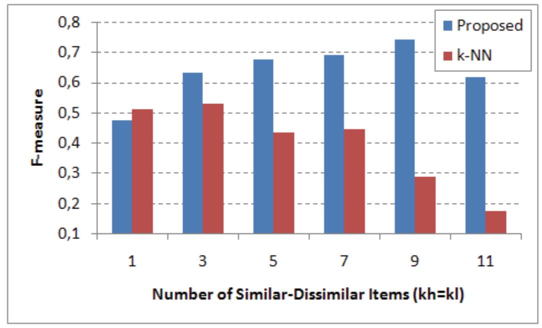

Fig. 4. Effect of similar (k_h) and dissimilar (k_l) items' quantity compared with the conventional k-NN.

Figure 5 displays the effect of similar items' (or top-matches) weight, w, on the retrieval performance. We observe that combining the information from the similar and dissimilar returns result in a better retrieval performance than using the individual contributions separately, most probably because in the image domain brains look highly alike despite the subtle differences present.

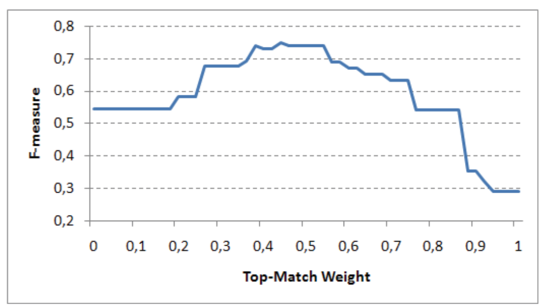

Fig. 5. Effect of similar item's (or top-matches) weight ($k_h = k_l = 5$).

A retrieval system's ability to perform equally at varying database sizes is referred to its scalability. In order to test this property, we evaluated our system with varying database sizes. As observed in Figure 6, after a quick-rise and some fluctuations the retrieval performance reaches to a relatively stable stage for database sizes above 60.

Current criteria used by physicians to diagnose dementia has good sensitivity at the expense of less optimal specificity or vice versa. A review of neuropathologically confirmed studies show that clinical diagnoses of common dementia types by experienced physicians have average sensitivities and specificities in the range of 22-93% and 48-100%, respectively [15]. Table 1 shows that our system has a sen-

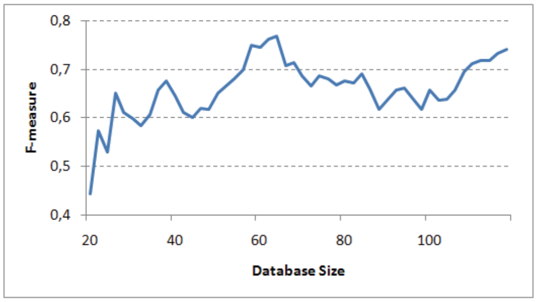

Fig. 6. Effect of database size.

Table 1. Confusion matrix.

	True Demented	True Control
Demented	29	17
Control	12	60
F-measure=.742	Sensitivity=.707	Specificity=.779

sitivity score comparable to those achieved by experts and a more balanced sensitivity-specificity combination.

4.4. Computational Complexity

Algorithms used in this work are implemented in C# and the average retrieval time per subject on an Intel Pentium processor (2.8 GHz) with 1G memory is measured as 2.31 seconds.

5. CONCLUSION AND FUTURE WORK

In this paper, we presented a novel and fast search and retrieval method for brain MR images for dementia diagnosis. The method utilizes a tailored version of histogram of oriented gradients as features, and realizes diagnosis using similar and dissimilar cases retrieved from a database. We evaluated its performance with comprehensive experiments on real data and observed that 1)the performance degrades with extreme quantization of the feature histograms and relatively small databases, and 2)combining the information from similar and dissimilar returns results in a more accurate diagnosis than using the individual contributions separately. When compared to the diagnostic performances in clinical practice, our method achieves a comparable sensitivity score and provides a more balanced sensitivity-specificity combination.

6. REFERENCES

[1] "2008 Alzheimer's disease facts and figures," Tech. Rep., Alzheimer's Association, www.alz.org, 2008.

[2] P. Scheltens, N. Fox, F. Barkhof, and C. De Carli, "Structural magnetic resonance imaging in the practical assessment of dementia: beyond exclusion.," *Lancet Neurology*, vol. 1, pp. 13–21, 2002.

[3] P. Scheltens, et al., "Qualitative assessment of cerebral atrophy on MRI: Inter- and intra-observer reproducibility in dementia and normal aging.," *European Neurology*, vol. 37, no. 2, pp. 95–99, 1997.

[4] J. Barnes, et al., "Differentiating AD from aging using semiautomated measurement of hippocampal atrophy rates.," *NeuroImage*, vol. 23, pp. 574–581, 2004.

[5] P. M. Thompson, et al., "Tracking alzheimer's disease.," *Ann. N.Y. Acad. Sci.*, vol. 1097, pp. 183–214, 2007.

[6] P. A. Freeborough and N. C. Fox, "MR image texture analysis applied to the diagnosis and tracking of alzheimer's disease.," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 475–479, 1998.

[7] S. Klöppel, et al., "Automatic classification of MR scans in alzheimer's disease.," *Brain*, vol. 131, pp. 681–689, 2008.

[8] Chih-Cheng Hsu, Wesley W. Chu, and Ricky K. Taira, "A knowledge-based approach for retrieving images by content," *IEEE Trans. Knowl. Data Eng.*, vol. 8, no. 4, pp. 522–532, 1996.

[9] H. K. Huang, et al., "Image-matching as a medical diagnostic support (DST) for brain diseases in children," *Comput. Med. Imag. Graph.*, vol. 29, no. 2-3, pp. 195–202, 2005.

[10] S. M. Smith, "Fast robust automated brain extraction," *Human Brain Map.*, vol. 17, no. 3, pp. 143–155, 2002.

[11] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE CVPR*, 2005, vol. 1, pp. 886–893.

[12] D. S. Marcus, et al., "Open access series of imaging studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *J. Cogn. Neurosci.*, vol. 19, no. 9, pp. 1498–1507, 2007.

[13] J. C. Morris, "CDR: current version and scoring rules.," *Neurology*, vol. 43, pp. 2412–2414, 1993.

[14] D. Unay, A. Ekin, and R. Jasinschi, "Local structure-based region-of-interest retrieval in brain MR images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 4, pp. 897–903, 2010.

[15] D. S. Knopman, et al., "Practice parameter: Diagnosis of dementia (an evidence-based review): Report of Quality Standards Subcommittee of the American Academy of Neurology," *Neurology*, vol. 56, no. 9, pp. 1143–1153, 2001.