MULTI-CHANNEL BRAIN ATROPHY PATTERN ANALYSIS IN NEUROIMAGING RETRIEVAL

Sidong Liu[1], *Weidong Cai*[1], *Lingfeng Wen*[2], *Dagan Feng*[1]

[1] Biomedical and Multimedia Information Technology (BMIT) Research Group, School of Information Technologies, University of Sydney, Australia

[2] Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, Sydney, Australia

ABSTRACT

The high throughput 3D neuroimaging datasets have posed great challenges for neuroimaging data retrieval. To achieve more accurate neuroimaging retrieval, various content-based retrieval approaches have been proposed. Recent studies showed a tendency of using the localized features extracted from a subset of the brain structures, instead of the global features extracted from whole brain. However, these studies relied heavily on specific pattern analysis techniques or clinical expertise. In this study we proposed a Multi-Channel pattern analysis approach to identify the most discriminative disease-sensitive brain structures for neurodegenerative disorders and thus to enhance neuroimaging retrieval. The preliminary results suggested that the proposed Multi-Channel pattern analysis approach could confidently identify the brain structures with atrophy and further improve the neuroimaging retrieval performance.

Index Terms **—** multi-channel analysis, neuroimaging retrieval

1. INTRODUCTION

Neuroimaging is becoming increasingly important in non-invasive diagnosis of neurological disorders, the assessment of therapy, and monitoring disorder progression with the advances in imaging instrumentation. More effective approaches for image management and retrieval are required to address the challenge of greater volume of imaging data. Retrieval of neurological images using image-related features has shown its capability in providing an alternative or complement to the traditional approaches of keyword-based searches [1, 2]. There are various methods or systems proposed in the field of the content-based neurological image retrieval [3-13]. The whole brain based features were often used in neuroimaging retrieval [3, 4]. However, the neurodegenerative disorders usually exhibit atrophy patterns in certain brain structures and exploring such patterns may facilitate the retrieval, *i.e.*, extracting features from these disease-specific brain structures, or brain structures of interest (BSOIs) instead of the whole brain.

There are two ways to select the BSOIs. The first way is based on the established pathological knowledge of the disorders. For example, *Ventricles* and *Hippocampus* were believed to be sensitive to Alzheimer's disease (AD) and used as AD biomarkers in [5]. Batty *et al.* [6] designed a knowledge-based mask comprising of 5 brain structures for retrieval of AD and mild cognitive impairment (MCI) cases. We have previously proposed a set of disease-oriented masks (DOMs) based on the clinical expertise and further adaptively modified them with t-maps [7]. The other way for BSOI selection is to apply the pattern analysis techniques on whole brain and then select the BSOIs based on certain significance metric, *e.g.*, a set of BSOIs were selected using cost sensitive variances [8] and prediction accuracies [9] by Support Vector Machine. Heckemann *et al.* [10] applied the two-one-sided-test (TOST) to derive the discriminative score for each region based on its *p-value*. Both two types of BSOI selection approaches have been reported to enhance the retrieval; however, the BSOIs used in these studies were not consistent with each other. This fact implied that complementary information from different approaches could be combined to improve the retrieval. In addition, since neurodegenerative disorders are very likely to be progressive, patients at severe stage usually exhibit a more spreading atrophy pattern than early stage. For example, MCI often represents a pre-symptomatic status of AD, conferring a high conversion rate of 16% to AD per annum [11]. A pattern of hypo-metabolism in *Hippocampus* and *Posterior Cingulate Cortex* was reported on MCI [12]; whereas whole brain was affected in more severe AD [13]. The atrophy pattern analysis is important for accurate neuroimaging retrieval and disorder differentiation, but it was less investigated in retrieval studies.

To faithfully select the disease-related brain regions, we presented a general method for pattern analysis integrating patterns extracted by various pattern analysis algorithms among different neurological disorders. Our method compensated the imperfections of individual algorithms, thus could describe the atrophy patterns with less bias and eventually enhance the neuroimaging retrieval.

*This work was supported in part by ARC grants.

2. METHODOLOGY

2.1. Neuroimaging Acquisition and Pre-processing

Data used in the preparation of this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database* [11]. Totally 369 subjects were randomly selected from the ADNI baseline cohort, and a T1 weighted 1.5T MRI image was acquired for each subject.

All 'raw' MRI data were converted to the ADNI format according to the ADNI MRI image correction protocol [14]. We then nonlinearly registered the MRI images to the ICBM_152 template [15], using the NREG algorithm [16] provided in the Image Registration Toolkit (IRTK) [17]. IRTK works in a coarse-to-fine manner with isotropic control point spaces ranging from 12mm to 1.5mm in 4 octaves. 17 MRI studies were excluded because intolerable morphometric distortions were observed in these registered MRI images. The participants include 85 AD subjects, 181 MCI subjects and 85 normal aging subjects.

Automatic labeling of 83 brain structures were achieved in the template space using the multi-atlas propagation with enhanced registration (MAPER) approach [8]. The atlas data required for MAPER comprising of 30 T1-weighted MRI images acquired from the National Society for Epilepsy at Chalfont, UK. The segmentation protocols were described in Hammers *et al.* [18]. The registration coefficients were transformed back to each MRI space to calculate the 83 BSOIs' volumes as the features for each subject.

2.2. Multi-Channel Atrophy Pattern Analysis

Brain atrophy pattern analysis could be carried out between different diagnosis group pairs using different analysis algorithms. We refer to each combination as a channel, and all channels are parallel to each other. Multifold single channel analysis could provide complementary information to each other. Therefore, in this study, we designed a multi-channel analysis algorithm to depict the neurodegenerative patterns of AD and MCI.

The ADNI dataset was first divided into AD, MCI and Normal Control (NC) groups based on the diagnoses. Three algorithms, Elastic Net, SVM and TOST, were then applied to analyze the neurodegenerative patterns in each group-pair. Finally, individual channels voted on all features/brain structures to select the most discriminative BSOIs.

We first designed three channels (*CH 1 – CH 3*) based on the classic TOST, which is probably the most commonly used test in neurodegerative pattern analysis [5, 10]. In a TOST, a given null hypothesis of equivalence, H_0, will be rejected when the *p-value* of the test stastic is smaller than a user-defined threshold. In this study, we set the significant value as 0.05. The *t-statistics* of feature variables whose *p-value* passed the significance test were outputted into one vector as the analysis result. TOST was carried out for AD/NC group pair, AD/MCI group pair, and MCI/NC group pair, respectively.

Another three channels (*CH 4 – CH 6*) were developed with Support Vector Machine (SVM). When SVM with the linear kernal is used to classify two groups of data, the separating hyperplane slopes, w, on each dimension could be used as the variable weights. Given a training set of instance-label pairs $(x_m, y_m), m = 1, \dots, M$ where $x_m \in \mathbb{R}^N$ and $y_m \in \{-1, 1\}^M$, the SVM solves the problem:

$$\hat{w} = \arg\min_w \frac{1}{2} w^T w + C \sum_{m=1}^{M} \xi_m \quad (1)$$
$$\text{s.t.} \quad y_m(w^T x_m + b) \geq 1 - \xi_m \; ; \xi_m \geq 0.$$

In this study, SVM based pattern analysis was performed in two steps. Firstly, we conducted a 10-fold cross validation on the three-category classification using the LIBSVM [19] to obtain the optimal estimate of cost parameter (C) for errors (ξ_m). When C was decided (C = 8 in this study), we then solved the problem in Eq. (1) using CVX, a package for specifying and solving convex programs [20]. Among all the regions, we selected the ones whose weights larger than the upper quartile of the entire weight distribution. SVM-based analysis was carried out for three channels, same as TOST.

We further set up four channels (*CH 7 – CH 10*) based on Elastic Net (EN). EN is a well established feature selection algorithm solving the problem:

$$\hat{\beta} = \arg\min_\beta \|y - X\beta\|^2 + \lambda_2 \|\beta\|^2 + \lambda_1 \|\beta\|_1 \quad (2)$$

where y is the response vector of M observations, X is the matrix of M feature vectors, $X = \{x_1 \dots x_m \dots x_M\}^T$, λ is a positive regulazation parameter, and β is the vector with same dimension as x_m. EN introduced l_1 and squared l_2 penalty of β, which could encourage grouping effect on feature variables and remove the limitation on the number of selected features. These characteritics make EN very suitable for studies with large number of observations and small number of features, or correlated features. The ADNI dataset is a typical case for EN, because the number of patients is much larger than the number of brain structures, and the brain structures were highly correlated. Another character of EN is that it can be used for multi-categorical feature selection, which is not supported by TOST and SVM. Therefore, we performed EN-based analysis not only on three group pairs, but also on the AD-MCI-NC triplet. So totally 10 channels were set up in parallel.

We finally integrated the outputs from individual channels for each feature variable as:

$$V := \{v_n, n = 1, \dots, N\} \quad (3)$$
$$v_n = \sum_{c=1}^{10} \mathrm{sgn}(v_{c,n}) \quad (4)$$

where $v_{c,n}$ is the output of the n^{th} feature variable in the c^{th} channel, and

$$\mathrm{sgn}(v_{c,n}) = \begin{cases} 1 & if\ v_{c,n} \neq 0 \\ 0 & if\ v_{c,n} = 0 \end{cases} \quad (5)$$

*For up-to-date information see www.adni-info.org.

This multi-channel scheme allowed the single channels to vote on each feature variable, and set higher values on the brain structures recognized by more channels. The votes of each region were highly correlated with the brain structure's discriminative power.

2.3. Supervised Pattern Optimization

Multi-channel pattern analysis could identify the disorder sensitive regions, but might result in an overlarge pattern involving too many brain structures. Certain less sensitive structures were undesired because they might be caused by biased group size or the imperfect analysis results, consequently might decrease the retrieval performance.

To optimize the pathological pattern derived from multi-channel analysis in neuroimaging retrieval, we proposed a thresholding-based pattern optimization approach, which tested a set of feature selection criteria and aimed to find the optimal subset of the brain structures associated with the multi-channel pattern. We first sorted the brain structures according to their votes based on our multi-channel analysis, and then carried out a series of retrieval experiments using subsets of the brain structures. The subsets were selected in a 'many to few' manner based on the structures' votes. We designed 6 experiments as shown in Figure 1. When the threshold became greater than 6, too few BSOIs were selected, thereby markedly decreasing the retrieval performance.

$BS := \{R_I | i = 1{:}83\}$ // The collection of all brain structures
$BSOI = \{\emptyset\}$ // The collection of selected structures
for $threshold = 1:6$ // Set the selection threshold
| **foreach** R_I **in** BS // Iterate each brain region
| | **if** $votes(R_I) > threshold$ **then** $BSOI = \{BSOI, R_I\}$
| **endforeach**
endfor
return $BSOI$ // Return the selected brain regions

Figure 1 The pseudo-code of the brain structure subset selection algorithm used in this study.

2.4. Performance Evaluation

We used the query by example paradigm in this study, and adopted the leave-one-out cross-validation on the whole database. The similarity between any two feature-vectors was calculated by the normalized mutual information.

We evaluated the performance using the mean average precision (MAP), *i.e.*:

$$\mathrm{MAP} = \frac{\sum_{q=1}^{Q} \sum_{k=1}^{K_q} (p_q(k) \cdot rel_q(k))/T_q}{Q} \quad (6)$$

where q is the index of the query, Q is the total number of queries, k is the rank in the sequence of retrieved image, K_q is the number of total retrieved image for the query, $p_q(k)$ is the precision at cut-off k in the result list, $rel_q(k)$ is the relevance score of the k^{th} retrieval result of the query, and T_q is the number of relevant images associated with the query. In this study, MAP was implemented slightly different from the classical way, because the group sizes of AD and MCI do not match. We assumed that there are at least 5 relevant cases exists for each query, so we set the $T_q = 5$ for all the queries. We also used elastic relevance criteria in this study, based on fact that MCI usually represents the transition state of NC to AD. Therefore, we set the relevance criteria as follows:

a. if the query is AD and the retrieval result is normal, or vise versa, then the relevance score is 0, *i.e.*:
$rel_{AD}(NC) = rel_{NC}(AD) = 0$;

b. if the retrieval result is from the same group of query, then the relevance score is 1, *i.e.*:
$rel_{AD}(AD) = rel_{MCI}(MCI) = rel_{NC}(NC) = 1$;

c. for all the other scenarios, we set the relevance score as 0.25 to model the transitional state of MCI in between AD and NC, *i.e.*:
$rel_{AD}(MCI) = rel_{MCI}(AD) = rel_{NC}(MCI) = rel_{MCI}(NC) = 0.25$.

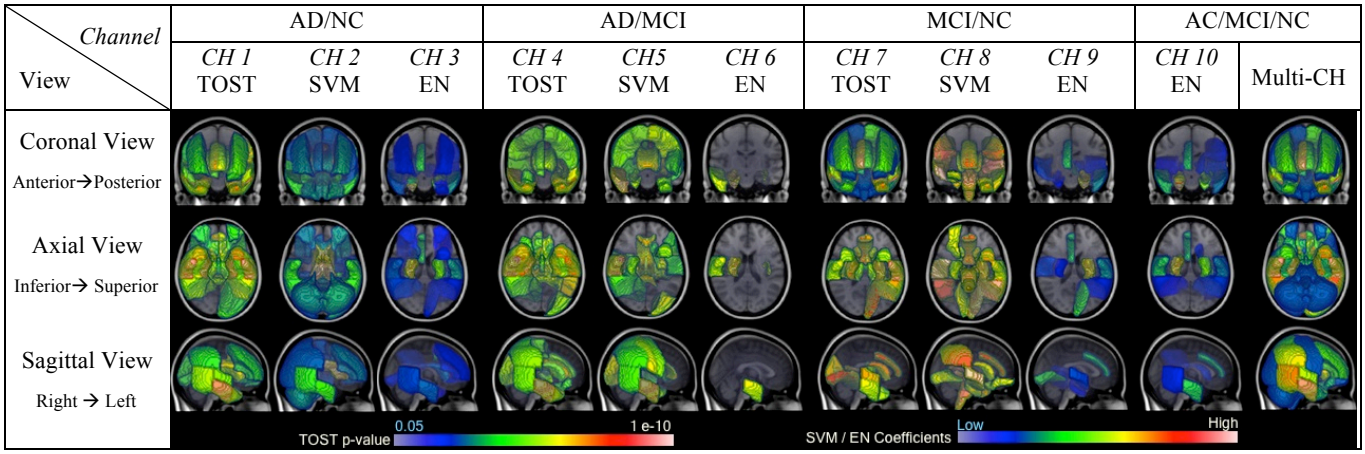

Figure 2 The derived brain atrophy patterns from individual single-channel analyses and multi-channel analysis in 3D views. The images were generated by 3D Slicer (Version 4.1) [21] and illustrated using the ICMB_152 template [15].

3. RESULTS

Figure 2 illustrates the derived patterns of each single channel analysis and our proposed multi-channel analysis. The differences between AD and NC were always larger than other group pairs. MCI tended to affect the *Parietal* and *Temporal Cortex*, whereas a more spreading atrophy pattern in partial *Frontal* and *Occipital Cortex* was observed in AD. Several subcortical structures were also identified as early detection biomarkers to distinguish MCI from NC, including *Parahippocampus*, *Cinguate Cortex*, and *Lingual Gyrus*. Figure 2 also shows the aggregated pattern derived from our proposed Multi-Channel approach. This near symmetrical pattern includes more brain structures compared to the individual single-channel patterns, and meanwhile the pattern's contrast was remarkably increased.

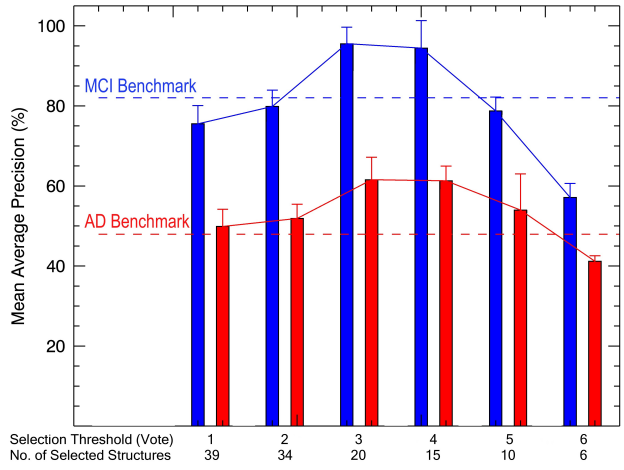

Figure 3 The retrieval performance using the derived patterns by the proposed approach, compared to benchmark results.

Figure 3 shows the retrieval performance with BSOI selection thresholds from 1 to 6. As the threshold incremented, fewer brain structures were retained. The *y*-axis indicates MAP±Standard deviation. Whole brain based features were used in the benchmark experiments, and their performances were shown as the dash lines. The retrieval using the derived patterns with appropriate structure selection threshold could achieve better results than the benchmark retrieval. However, retrieval based on too few or too many BSOIs could lead to lower performance because of information loss or noise gain. For both AD and MCI, the MAPs of different thresholds formed a convex curve and peaked when the selection threshold was set as 3 and 20 BSOIs were selected. The highest MAP for AD was 61.6% (13.6% higher than AD benchmark); and the best retrieval rate for MCI is 95.5% (13.5% higher than MCI benchmark). The selected BSOIs, as shown in Figure 1 with higher values, mostly spread across the *Temporal* and *Parietal Lobes*, partial *Cingulate Cortex*, *Ventricles* and *Amygdala*.

4. CONCLUSION

In this work we presented a multi-channel brain atrophy pattern analysis algorithm for neuroimaging retrieval. The main advantage of our algorithm is its capability to learn atrophy patterns of progressive neurological disorders and also overcome the drawbacks of individual analysis techniques. We further optimized the pattern by a supervised thresholding method, and consequently 20 brain structures were identified as the most discriminative biomarkers for AD and MCI.

5. REFERENCES

[1] H. Muller, N. Michoux, *et al.,* "A Review of Content-based Image Retrieval Systems in Medical Applications – Clinical Benefits and Future Directions," *Int. J. Med. Info.*, vol. 73, pp.1-23, 2004.

[2] W. Cai, J. Kim and D. Feng, "Chapter 4 – Content-based Medical Image Retrieval", in *Biomed. Info. Tech.*, pp. 83-113, 2008.

[3] S. Liu, W. Cai, *et al.*, "Generalized Regional Disorder-Sensitive-Weighting Scheme for 3D Neuroimaging Retrieval", *EMBC2011*, pp. 7009-7012, 2011.

[4] S Liu, W Cai, *et al.*, "Multiscale and Multiorientation Feature Extraction with Degenerative Patterns for 3D Neuroimaging Retrieval", *ICIP2012*, 2012.

[5] Y. Wang, Y. Song, *et al.*, " Surface-based TBM Boosts Power to Detect Disease Effects on the Brain: An N = 804 ADNI Study," *Neuroimaging*, vol. 56, pp. 1993-2010, 2011.

[6] S. Batty, *et al.*, "Prototype System for Semantic Retrieval of Neurological PET Images", *Int. Conf. Med. Image. Info.*, pp.179-188, 2008.

[7] W. Cai, S. Liu, *et al.*, "3D Neurological Image Retrieval with Localized Pathology-Centric CMRGLC Patterns", *ICIP2010*, pp. 3201-3204, 2010.

[8] D. Zhang and D. Shen, "MultiCost: Multi-stage Cost-Sensitive Classification of Alzheimer's Disease," *MLMI2011*, pp. 344-351, 2011.

[9] P. M. Rasmussen, *et al.*, "Model Sparsity and Brain Pattern Interpretation of Classification Models in Neuroimaging, " *Pattern Recognition*, vol. 45, pp. 2085-2100, 2012.

[10] R. A. Heckemann, *et al.*, "Automatic Morphometry in Alzheimer's Disease and Mild Cognitive Impairment," *Neuroimage*, vol. 56, pp. 2024-2037, 2011.

[11] W. J. Jagust, D. Bandy, *et al.*, "The Alzheimer's Disease Neuroimaging Initiative Positron Emission Tomography Core," *Alzheimer's & Dementia*, vol. 6, pp. 221-229, 2010.

[12] L. Mosconi, W.H. Herholz, *et al.*, "Multicenter Standard 18F FDG PET Diagnosis of Mild Cognitive Impairment, Alzheimer's Disease," J. Nuclear Med., vol. 49, pp. 390-398, 2008.

[13] J. Langbaum, K. Chen, *et al.*, "Categorical and Correlation Analyses of Baseline FDG-PET Images from the ADNI", *Neuroimaging*, vol. 1, issue 45, no. 4, pp. 1107-1116, May 2009.

[14] C. R. Jack, M. A. Bernstein, *et al.*, "The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI Methods," *J. Magnetic Resonance Imaging,* vol. 27, issue 4, pp. 685 – 691, 2008.

[15] J. Mazziotta, *et al.* "A Probabilistic Atlas and Reference System for the Human Brain: International Consortium for Brain Mapping (ICBM)," *Phil. Trans. Royal Soc. B Biol. Sci.*, vol. 356, no. 1412, pp. 1293-1322, 2001.

[16] J. A. Schnabel, *et al.*, "A Generic Framework for Non-rigid Registration based on Non-uniform Multi-level Free-form Deformations," *MICCAI2001*, pp. 573-581, Utrecht, NL, October 2001.

[17] D. Reuckert, L.I. Sonoda, *et al.*, "Nonrigid Registrationusing Free-Form Deformations: Application to Breast MR Images," *IEEE Trans. Med. Imaging*, vol. 18, no. 8, pp. 712-721, 1999.

[18] A. Hammers, *et al.*, "Three-Dimensional Maximum Probability Atlas of the Human Brain, with Particular Reference to the Temporal Lobe, " *Human Brain Mapping*, vol. 19, no. 4, pp. 224-247, 2003.

[19] C. C. Chang, and C. J. Lin, "LIBSVM : a library for support vector machines," *ACM Trans. on Intel. Sys. Tech*, vol. 2, no. 27, pp. 1:27, 2011.

[20] CVX Research, Inc. CVX: Matlab software for disciplined convex programming, version 2.0. April 2011, http://cvxr.com/cvx.

[21] S. Pieper, *et al.*, "The NA-MIC Kit: ITK, VTK, Pipelines, Grids and 3D Slicer as an Open Platform for the Medical Image Computing Community," *ISBI2006*, vol.1, pp. 698-701, 2006.