Classification of Alzheimer's disease using structural MRI

Chester V. Dolph, Manar D. Samad, Khan M. Iftekharuddin

Vision Lab, Old Dominion University, Norfolk, VA 23529

Abstract. We propose an Alzheimer's disease (AD) classification method that uses selected features in segmented brain tissue from individual MRI slices. The objective of our method is to aid physicians in classifying AD progression, particularly for subjects with Mild Cognitive Impairment. The proposed method captures atrophy of regions of the brain without extracting specific regions such as the hippocampus. All subjects are registered to an atlas designed for AD research obtained from Laboratory of Neuro Imaging (LONI) at University of Southern California (USC). Each slice is segmented into grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF) to obtain the number of instances of these tissues per slice. The features include the number of instances of a segmented tissue in a given slice and the ratio of WM to CSF. The proposed method uses existing tools for skull stripping, registration, segmentation, and feature classification. A 10-fold cross-validation on CADDementia training data of 30 MRI samples yields 80% classification accuracy in classifying the 3 different cognitive states such as Cognitive Normal (CN), Mild Cognitive Impairment (MCI), and Alzheimer's disease (AD). Our results show that all AD and CN patients are classified correctly, however, subjects with MCI are misclassified mainly as the CN class. Although a larger training data can enhance the classification result, we use this limited training data to obtain our test results for 354 patients for this challenge.

1 Introduction

The progression of AD is of great interest in medical research as 1 in 3 seniors in the United States dies with dementia [1]. Based on signs and symptoms, physicians usually track AD using the Clinical Dementia Rating (CDR) system. Using CDR, subjects are classified in three states such as CN, MCI, and AD. The progression of AD can be characterized by atrophy of GM and WM along with expansion of CSF volume. These structural changes in brain facilitate the distinction of an AD brain from a CN brain; however, the distinction between MCI and CN is subtle [2]. Previous studies on classifying AD from MRI data include features based on voxel intensity, volume, regional, and thickness of subcortical areas [3]. Considering individual voxels of entire volume of MRI data is computationally expensive. On the other hand, focusing on a particular region of the brain may miss important structural changes happening in other regions of the brain.

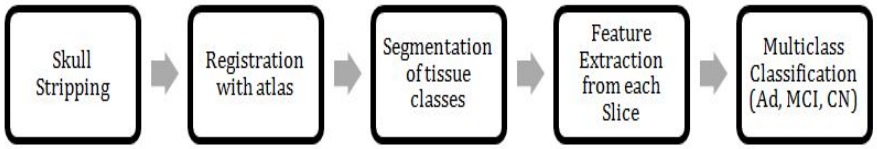

Fig. 1. Proposed algorithm pipeline.

Other methods include segmentation of the hippocampus [4]. For the MCI class, where the changes are subtle, considering only a particular region of the brain may not be helpful in classification. Rather than using individual voxels or a particular brain region, we consider total number of pixels per slice from the entire MRI volume for the purpose of feature extraction and classification of the 3 different cognitive states. From each slice, the proposed method is based on the total number of pixels per slice classified as GM, WM, and CSF and their proportional relationships. The benefit of this approach is dimensionality reduction where the structural changes in the entire MRI data are represented in a feature dimension equal to the number of slices.

Since AD progresses differently among subjects, our approach captures features from selected MRI slices while accounting for tissue atrophy. Starting with all the slices to probe for structural changes in different tissues, this work applies a trial and error method to select the slices that best represent the structural changes for the best performance of 3-class classification. The search for the best slices starts from the mid region of the brain where the slices encompass most of the brain tissues.

2 Methods

2.1 Training data

The training data used for this work is obtained from the non-uniformity corrected data from the CADDementia website [5]. This data set consists of 30 subjects structural MRI labeled AD, MCI, or CN. This training data set contains MR image data labeled for 9 AD, 9 MCI, and 12 CN patient classes.

2.2 Algorithm description

The data processing pipeline is shown in Figure 1. The pipeline consists of five major steps such as skull stripping, registration, segmentation, feature extraction from slices, and classification. The steps of the pipeline are discussed in details in the following paragraphs.

Data pre-processing Prior to feature extraction, the following pre-processing steps are employed within the pipeline.

Skull stripping The MRI from each subject is skull stripped using the Brainsuite software tool [6]. In general, default parameters are used in the skull stripping process. However, if any scalp remains after stripping with default parameters, the diffusion and iteration parameters are adjusted as necessary to remove the remaining skull.

Registration Subject MRI scans are registered to the Alzheimer's disease atlas obtained from the LONI at USC using Deformable Registration via Attribute Matching and Mutual-Saliency Weighting (DRAMMS) [7].

Segmentation The brain volumes are segmented into GM, WM, and CSF using the Brainsuite classification tool [8].

2.3 Feature Extraction

Feature extraction: The features for our classifier are based on the segmented tissues. GM and WM atrophy is expected of the AD class and to a less extent the MCI class, while CSF volume simultaneously expands . We sum each segmented class for each MRI slice to create features that capture atrophy of that specific segmented brain tissue class. The sum represents the total number of instances that a segmented tissue is classified as GM, WM, or CSF. For example, a slice could be described as having 100 pixels, which might be further decomposed as 40 GM pixels, 35 WM pixels, and 25 CSF pixels after the tissue segmentation step. Since we are expecting fewer instances of GM and slightly more instances of CSF for a typical AD subject compared to a CN subject, the ratio of total pixels for GM (given as slice GM) to total number of pixels for CSF (slice CSF) should yield a larger number than the corresponding ratio of a cognitively normal subject. We propose the features given as,

$$\frac{slice\ WM}{slice\ CSF}, \quad \frac{slice\ GM}{slice\ CSF}, \quad slice\ CSF, \quad slice\ WM, \quad slice\ GM, \tag{1}$$

where slice CSF represents total number of pixel for tissue CSF in an MRI slice and so on. A feature selection step is involved in order to select the MRI slices that yield the best classification results. For 3-class classification, different multi-class classifiers such as support vector machine SVM (with both Polynomial, and RBF kernels) and random forest are evaluated for accuracy on the training data. A 10-fold cross-validation is performed to find the best classifier and the feature using the training data set. We use the combination of best performing feature and classifier to classify the 354 test MRI samples for this challenge.

3 Results and discussions

3.1 Algorithm performance

Computation of skull stripping and segmentation takes approximately 1 minute per subject, which is mainly the time to navigate on the graphical user (GUI) interface of Brainsuite tool. Registration time varies for subjects from about 15 to 20 minutes each for each volume. The registration is completed serially on an i7 processor with 8 gigabytes of RAM. Increasing the memory capacity of the machine has the potential of speedup, as currently, there is excessive memory swap. Analysis of the segmented tissues is performed serially in Matlab on a MacBook with i5 processor and 8 gigabytes of RAM. This process is parallel, thus it has potential for significant speedup. Since the proposed feature dimension is only 46 per subject, it took only several seconds to classify the test data with 354 subjects.

The proposed algorithm is semi-automatic. The first few pre-processing steps such as skull stripping and segmentation need manual intervention only to adjust some parameters using the Brainsuite software tool. The Brainsuite GUI requires only a few parameter modifications per subject for skull stripping. These adjustments typically require 30 seconds of time, however, roughly 45 subjects required 15 minutes of parameter selection to remove the entire scalp. The subject registration is batch scripted in DRAMMS and thus completely automatic. The remaining steps including per slice analysis of the segmented tissues, feature extraction, and classification are all automated as well. Overall, the proposed pipeline takes about 30 minutes per subject.

4 Classification performance

An MR image of a subject's brain is a 3D volumetric data consisting of 128 slices. Following Eq. (1), each feature value represents the ratio of a pair of tissue content in each slice. From pattern recognition perspective, not all slices may contain discriminatory information that may be useful for 3-class classification. There may be slices with redundant or irrelevant information, which may introduce ambiguity in different class boundaries under classification. Following a trial and error step, 46 slices from slice index 25 to 70 for each subject MR data are found to be the most effective set of slices for the proposed 3-class classification. Figure 2 shows the range of the slices in the 3D volume. Fig. 2 shows that the slices containing relevant subcortical areas such as hippocampus plays major role in the proposed classification scheme.

The features proposed in Section 2.2.2 are extracted and used from these 46 slices. For each type of feature, there are 46 attributes per subject. Therefore, for 30 subjects in the given training data set, the training data dimension is 30 by 46. 3-class classification is performed by training 3 possible binary classifiers. Results from these binary classifiers are combined to yield the final multi-class classification result. A 10-fold cross-validation is performed to evaluate the performance of our proposed features in classifying 3-class MR images of brain.

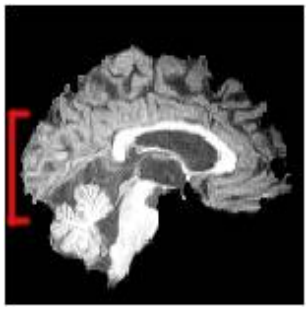

Fig. 2. Location of the cross-sectional slices used for classification.

Before testing a fold of data, the binary classifiers are trained by the remaining 9-fold data. The final 10-fold classification accuracy is obtained by averaging the accuracy from 10 individual folds of test data.

Table 1 shows the confusion matrix after 10-fold cross-validation using WM/CSF feature and SVM with RBF kernel (SVM-RBF) classifier. Note that all the AD and CN classes are correctly classified. However, MCI class is mostly confused with CN class. The 10-fold cross-validation result shows 80% classification accuracy wherein 24 out of 30 samples are correctly classified in multi-class classification. Table 2 shows area under the ROC curve for 3 different classes and different classifiers . For classification, an SVM-RBF classifier is found to yield the best

Classified As –>	AD	MCI	CN
AD	9	0	0
MCI	1	3	5
CN	0	0	12

Table 1. Confusion matrix for training data set after 10-fold cross-validation.

classification accuracy. Out of five different features, we find that (WM/CSF) feature performs the best in classifying 3 classes. Therefore, we consider the same SVM-RBF classifier and WM/CSF feature to classify the test data set.

The absolute proportion of WM, GM, and CSF yields poor classification accuracy. This is understandable, since different subjects are likely to have different proportion of tissues regardless of their cognitive status. Therefore, to have a representative feature for a cognitive status regardless of the subject, the ratio of different tissues within a subject can be a subject independent attribute. Since subjects with cognitive impairment suffer atrophy in WM or GM with expansion in CSF region, a ratio between WM or GM with CSF can be representative feature for such classification. This is supported by our results shown above,

Classifier	AD	MCI	CN	Average
SVM-RBF	0.976	0.762	0.861	0.866
SVM-Poly	0.881	0.574	0.88	0.788
Random Forest	0868	0.624	0.861	0.792

Table 2. Area under the ROC curve for each class

which reveals a significant improvement in the classification performance using (WM/CSF) feature.

5 Conclusion

We eagerly await the classification accuracy of the testing data set, so that conclusions can be made on the performance of this proposed approach. The classification accuracy also depends on the success and consistency in pre-processing steps as well as on the training data size. Furthermore, the relatively poor classification between MCI and CN is due to the subtle differences between these two groups. We believe that a larger training data set will yield an improved performance in the classifying the test samples of MR images from 354 subjects.

Acknowledgments

This work is partially supported through grants funding from NIH (R15CA115464). Authors would like to thank Dr. Christos Davatzikos for his kind assistance on the DRAMMS tool.

References

1. Alzheimer's Association, 2013. Alzheimer's News 3/19/13, http://www.alz.org/news_and_events_facts_and_figures_report.asp.
2. Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehricy, S., Habert, M.O., Chupin, M., Benali, H., Colliot, O., 2011. Automatic classification of patients with Alzheimers disease from structural MRI: A comparison of ten methods using the ADNI database. Neuroimage 56, 766781.
3. Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox, N.C., Jack Jr., C.R., Ashburner, J., Frackowiak, R.S.J., 2008. Automatic classification of MR scans in Alzheimer's disease. Brain 131 (3), 681689.
4. Chupin, M., Hammers, A., Liu, R.S., Colliot, O., Burdett, J., Bardinet, E., Duncan, J.S., Garnero, L., Lemieux, L., 2009a. Automatic segmentation of the hippocampus and the amygdala driven by hybrid constraints: method and validation. Neuroimage 46 (3), 749 761.
5. A standardized evaluation framework for Computer-Aided Diagnosis of Dementia based on structural MRI data, http://caddementia.grand-challenge.org/.

6. Shattuck DW and Leahy RM (2002) BrainSuite: An Automated Cortical Surface Identification Tool Medical Image Analysis 8(2):129-142.
7. DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting. Yangming Ou, Aristeidis Sotiras, Nikos Paragios, Christos Davatzikos. Medical Image Analysis, 15(4): 622-639, 2011.
8. Shattuck DW, Sandor-Leahy SR, Schaper KA, Rottenberg DA, and Leahy RM (2001) Magnetic Resonance Image Tissue Classification Using a Partial Volume Model NeuroImage 13(5):856-876.