

# A graphical model approach to ATLAS-free mining of MRI images\*

Chris S. Magnano<sup>†</sup>    Ameet Soni<sup>‡</sup>    Sriraam Natarajan<sup>§</sup>    Gautam Kunapuli<sup>¶</sup>

## Abstract

Improvements in medical imaging techniques have provided clinicians the ability to obtain detailed brain images of patients at lower costs. This increased availability of rich data opens up new avenues of research that promise better understanding of common brain ailments such as Alzheimer’s Disease and dementia. Improved data mining techniques, however, are required to leverage these new data sets to identify intermediate disease states (e.g., mild cognitive impairment) and perform early diagnosis.

We propose a graphical model framework based on conditional random fields (CRFs) to mine MRI brain images. As a proof-of-concept, we apply CRFs to the problem of brain tissue segmentation. Experimental results show robust and accurate performance on tissue segmentation comparable to other state-of-the-art segmentation methods. In addition, results show that our algorithm generalizes well across data sets and is less susceptible to outliers. Our method relies on minimal prior knowledge unlike atlas-based techniques, which assume images map to a normal template. Our results show that CRFs are a promising model for tissue segmentation, as well as other MRI data mining problems such as anatomical segmentation and disease diagnosis where atlas assumptions are unreliable in abnormal brain images.

**Keywords:** Graphical models; image segmentation; brain images

## 1 Introduction

Magnetic resonance imaging (MRI) is a neuroimaging method that can be used for visualization of brain anatomy with a high degree of spatial resolution and contrast between brain tissue types. Structural MRI methods have been used to identify regional volumetric changes in brain areas known to be associated with diseases such as Alzheimer’s (AD), demonstrating the utility of such methods for studying diseases [26, 29]. In the context of Alzheimer’s, structural MRI has identified associated cross-sectional differences and longitudinal changes in volume and size of specific brain regions, such as the hippocampus and entorhinal cortex, as well as regional alterations in gray matter, white matter and cerebrospinal fluid on a voxel-by-voxel basis [29].

We focus on mining and extracting useful informa-

tion from these structural MRI images. Many image techniques such as thresholding, region growing, statistical models, active-control methods and clustering have been previously used for medical imaging; for an in-depth review of MRI mining methods, please refer to Balafar et al. [4]. As pointed out in Balafar et al., thresholding methods by themselves are not ideal since distribution of intensities in medical images is quite complex. Probabilistic classification methods [7, 25] appear to be best suited for this complex task due to their ability to handle noise and ease of implementation.

At the other end of the spectrum are atlas-based segmentation methods that parcellate the MRI data into different anatomically relevant regions. For example, the Automated Anatomic Label (AAL) atlas<sup>1</sup> as implemented by *WFU PickAtlas* [20], divides the MRI image into 116 clinically important regions. Ideally, effective MRI-based diagnosis and identification relies on a segmentation approach to be (1) robust to noise, (2) able to handle large variances in brain intensities and (3) subject- and disease-specific.

While atlas-based brain warping works well for normal brains, it fails to capture the morphological changes that could result from brain diseases such as tumors, Alzheimer’s, etc., [1, 2]. This is because atlas-based segmentation is not subject- or disease-specific; quite the contrary, all brains are segmented into the same set of regions, and assume normal characteristics for each region (i.e., that there is a one-to-one mapping to a template). This results in a decrease in shape variability i.e., it manifests in an inability to capture local shape and feature information, which is essential for discriminating between brains for diagnosis of diseases.

This issue is addressed in our work: our motivation is to find a discriminative machine-learning-based approach to segment brain images and detect abnormalities that could *aid in classification of diseases from images*. Recent work [22] on predicting the incidence of AD using EM-based segmentation combined with an ensemble method demonstrates that such an approach produces superior diagnostic performance when applied to structural MRI data. We further motivate our approach with another example that previews the work

\*Supported by the Howard Hughes Medical Institute (HHMI)

<sup>†</sup>Swarthmore College, Computer Science Department, cmagnan1@swarthmore.edu

<sup>‡</sup>Ibid., soni@cs.swarthmore.edu

<sup>§</sup>Indiana University, School of Informatics and Computing, natarasr@indiana.edu

<sup>¶</sup>UtopiaCompression Corporation, Los Angeles, CA, gautam@utopiacompression.com

<sup>1</sup><http://prefrontal.org/blog/2008/05/brain-art-aal-patchwork>

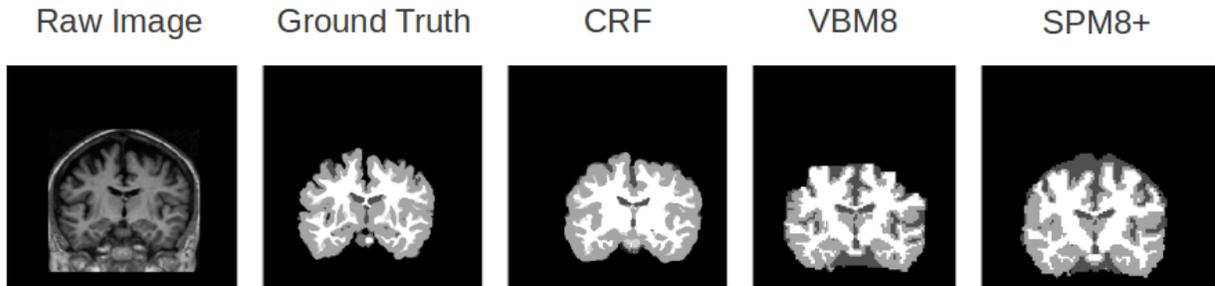


Figure 1: Illustration of tissue segmentation task using one 2D slice from a sample 3D MRI. From left to right: *Raw Image* – the intensity values obtained from MR (i.e., the algorithm input); *Ground truth* – a manually segmented image classifying each voxel as white matter (white), grey matter (light grey), or cerebral spinal fluid (dark grey); *CRF* – the proposed algorithm in this paper; *SPM8+*, *VBM8* – atlas-based baseline techniques representing current state-of-the-art methods in medical imaging.

presented here, and highlights the need for moving away from atlas-based methods. Consider Figure 1, which shows the results of using our probabilistic discriminative method against the standard method based on a Gaussian-mixture model (SPM)<sup>2</sup> [3, 9] and voxel-based morphometry (VBM) [2] for identifying important features (regions) when predicting Alzheimer’s disease. First, it is evident that our approach more closely resembles the ground truth, and thus is more effective in isolating the gray matter and cerebrospinal fluid intensities from the brain images compared to SPM and VBM. Second, in the raw image in Figure 1, there is a brain deformity in upper-right region. Mapping to a template atlas, however, assumes that each voxel in this deformed region maps to a voxel in a normal brain, even though it is more likely that this tissue is simply missing. Thus, SPM rounds out the image and erroneously fills it in with cerebrospinal fluid intensities while VBM transitions incorrectly from white matter to background.

While several diverse paradigms exist for image segmentation, we focus on probabilistic models, as they have been used successfully in many image segmentations tasks. For example, Friedman and Russell [11] utilized probabilistic models to detect objects in motion in video images. In particular, Markov Random Fields (MRFs) [12], a type of probabilistic graphical approach, have been applied to a wide variety of tasks including texture analysis and image restoration (a more thorough sampling of applications of MRFs for image

analysis can be found in Li [18]). MRFs have also been successfully applied to brain MRI segmentation [14, 30] and tissue classification [28]. MRFs, however, are a generative probabilistic model. That is, when our goal is to discriminate the tissue type of each voxel in a brain image, MRFs generate all the configurations of possible images as well as underlying tissue types, which can become extraordinarily inefficient.

Building off this success, recent years have seen the emergence of Conditional Random Fields (CRFs) [16, 27], which are a discriminative variant of MRFs; they have added the ability to model complex local dependencies in image-mining tasks, including labeling image regions on multiple scales [13] and object recognition [24]. In particular, Lee et al., [17] applied CRFs to a set of alignment-based features to perform brain-tumor segmentation for radiation-therapy target planning. CRFs have shown superior performance across many tasks as they directly optimize the classification task at hand (e.g., the possible configurations of tissue types given a fixed set of image intensities); for instance, the work of Kumar and Hebert [15] shows that CRFs outperform MRFs at modeling spatial dependencies across a diverse set of natural images. Inspired by these successes, we were motivated to develop a novel CRF-based segmentation approach for MRIs.

The key contributions of the paper are as follows: we (1) propose the first of its kind, fully-CRF-based framework for structural-MRI-image analysis (in particular for identification of relevant regions), and apply this to the task of volumetric segmentation for 3-dimensional data; (2) demonstrate the impact of the state-of-the-art algorithms for solving CRFs on MRI image analysis; (3)

<sup>2</sup>In this paper, we use **SPM8+** to refer to the *New Segment SPM8* algorithm, which is in beta but shows superior performance to to *Segment SPM8* [9]

apply our approach to standard brain image repository data sets; and, (4) show that our approach generalizes across data sets, an important feature for developing efficient disease- and subject-specific approaches.

This paper is organized as follows: after reviewing the general CRF probabilistic model, we will present our methodology in the following section. Next, we present the results of comparing our method to 3 other standard methods for white matter(WM) tissue classification and grey matter(GM) tissue classification. We conclude by discussing possible future work directions.

## 2 Background

Probabilistic graphical models encompass a set of expressive techniques for modeling the structural dependencies between random variables using a graph. In this paper, we examine a specific graphical model known as Conditional Random Fields (CRFs) [16] (for a thorough overview of CRFs, see Sutton and McCallum [27]).

CRFs model a probability distribution over two sets of random variables  $\mathbf{x}$  and  $\mathbf{y}$ . Here,  $\mathbf{x}$  represents the observed data, that is, the set of random variables representing evidence. For example, in a standard part-of-speech (POS) task,  $\mathbf{x}$  can represent the words in a sentence. The variables,  $\mathbf{y}$ , represent the underlying phenomena to analyze; this is known as the set of latent, or hidden, random variables in the posed problem. In the sample POS task,  $y_i \in \mathbf{y}$  encompasses the underlying part-of-speech tag for word  $i$  in the sequence. The data mining task is to find an assignment (or label)  $\hat{y}_i$  for each random variable  $y_i$  that maximizes the conditional probability of  $\mathbf{y}$  given the observed data,  $\mathbf{x}$ :

$$(2.1) \quad \hat{\mathbf{y}} = \arg \max_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}).$$

We define a graph  $G = (V, E)$  where each variable indexed in  $\mathbf{y}$  (i.e.,  $y_i \in \mathbf{y}$ ) corresponds to one vertex  $i \in V$ . Each edge  $e \in E$  represents a pairwise connection between two hidden variables and encodes a dependence between the two variables. In our figures, we also add the observed variables  $\mathbf{x}$  to the graph, and add corresponding edges connecting the evidence to the latent variables. CRFs assume the following Markov property on the graph, which is a key assumption of CRFs: each hidden variable  $y_i \in \mathbf{y}$ , given evidence  $\mathbf{x}$  and the value of its neighbors  $\mathbf{N}(y_i)$ , is conditionally independent of all other hidden variables:

$$(2.2) \quad P(y_i | \mathbf{x}, \mathbf{y}/y_i) = P(y_i | \mathbf{x}, \mathbf{N}(y_i)).$$

Notice that the model is *globally conditioned on all the evidence*. This is in contrast to MRFs which add constraints requiring the evidence variable  $x_i$  being conditionally independent given the label  $y_i$ .

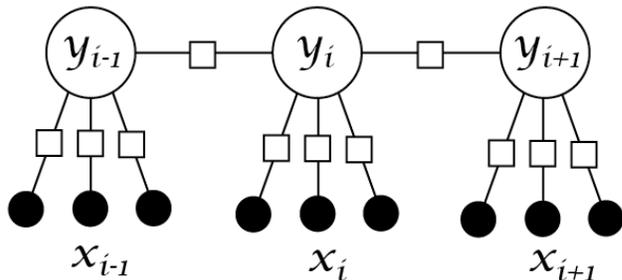


Figure 2: An example of conditional random field using a factor graph representation. The set of white nodes (e.g.,  $y_i$ ) represents hidden variables that we wish to infer. The black nodes (e.g.,  $\mathbf{x}_i$ ) represent sets of evidence variables. Each edge contains an edge potential (shown as white squares) that represent the encoded dependence between variables. Note that while CRFs do not require an imposed structure on the evidence data, our representation utilizes a local neighborhood of data for each hidden variable.

Figure 2 shows a sample CRF graph model using a factor graph representation. While the CRF can be conditioned on all evidence data,  $\mathbf{x}$ , we simplify to only consider a local neighborhood of evidence (e.g., the 3 words surrounding position  $i$  in POS tagging). Thus, each hidden variable  $y_i$  has a corresponding set of evidence variables  $\mathbf{x}_i$ . Using the factor graph representation, we break our model into two sets of binary feature functions:  $f(y_i, \mathbf{x}_i)$  (i.e., the connection between the label of a hidden variable and its evidence data) and  $f(y_i, y_j, \mathbf{x}_i)$  where  $e \equiv (i, j) \in E$  (i.e., the connection between the label at  $y_i$  and  $y_j$ ). To simplify notation, we will generalize the first set of features (i.e., add  $y_j$  as a parameter that is ignored). This leads to the following formulation of a CRF, whose probability distribution of the latent variables can be conditioned on the evidence as

$$(2.3) \quad p(\mathbf{y} | \mathbf{x}) = \frac{1}{Z(\mathbf{x})} \exp \left\{ \sum_{(i,j) \in E, k} \lambda_k f_k(y_i, y_j, \mathbf{x}_i) \right\},$$

and  $Z(\mathbf{x})$  is a normalization factor,

$$(2.4) \quad Z(\mathbf{x}) = \sum_{\mathbf{y}} \exp \left\{ \sum_{(i,j) \in E, k} \lambda_k f_k(y_i, y_j, \mathbf{x}_i) \right\}.$$

The variable  $k$  indexes each feature function in our set and  $\lambda_k$  is the weight learned for each feature  $f_k$ .

## 3 CRFs for MRI Image analysis

In this section, we outline our proposed method for identifying relevant regions from MRI images. Figure 3

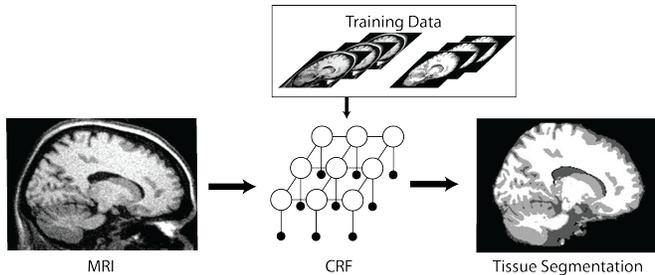


Figure 3: Pipeline of the proposed algorithm. The input is a CRF with a pre-defined structure, whose parameters are learned given training data. This model is then used for isolating tissues in a new MRI image.

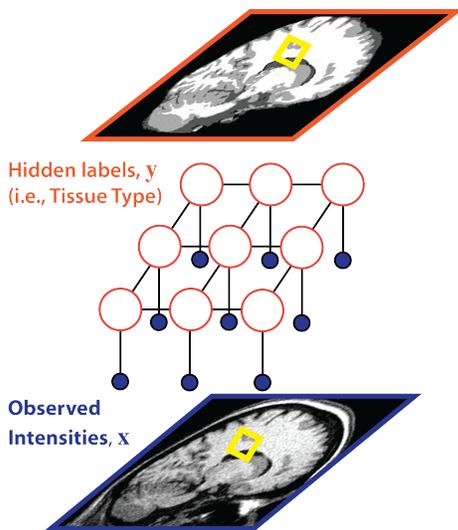


Figure 4: Structure of the CRF used in this work. Each voxel’s tissue type from the 3D image becomes a hidden node in the CRF. Each hidden node has 26 neighbors and 3 observations associated with it. MRI images are the observed variables in the model. The yellow box indicates that the shown CRF corresponds to representing the  $3 \times 3$  area of the 2D image slice.

presents the pipeline of the approach. The input to the algorithm is a CRF whose structure is predefined and a set of training examples in the form of MRI images and their corresponding tissue segmentations. Tissue segmentation here refers to classifying each tissue (voxel) as one of gray matter (GM), white matter (WM) and cerebral spinal fluid (CSF). Each MRI high (1.5mm) resolution image has about 3 million voxels. The CRF that is constructed would correspondingly have about 3 million hidden nodes (variables) each corresponding to a voxel in the MRI image.

Figure 4 provides more details about how the CRF

model corresponds to the input MRI. In our CRF model, each voxel has a corresponding hidden node  $y_i$  and can take one of three values,  $y_i = \{WM, GM, CSF\}$  representing white matter, grey matter and cerebrospinal fluid. We chose to model each voxel as being connected to 26 neighboring voxels (i.e., a  $3 \times 3 \times 3$  neighborhood around the voxel) with the result that each hidden node is connected to 26 other hidden nodes. Observations include the image intensity of each voxel, which is a function of the density of the underlying tissue type. We also included the mean intensities of the neighboring voxels and the Euclidean distance of the voxel to the center of the image as additional features. Associated with each voxel (and hidden variable)  $Y_i$  is a set of observations  $\mathbf{x}_i$  that include its own intensity value, the average intensities of the 26 neighbors and the position of the voxel.

Note that estimating the conditional distribution corresponds to estimating  $\lambda$  for the feature set. Also note that the structure CRF for each image could be different due to different brain sizes, but they all share the same set of parameters due to the features being the same for each node in the CRF. Put another way, this means that while the CRF graph-model can be of different sizes due to differences in the brain sizes, the set of parameters to be learned for each of these CRFs is the same. This is evident in Figure 4, which shows that each voxel type is the hidden node for the CRF, and the observed intensities along with the position become the observations for the CRF. While we show at most 8 neighbors (2D slice) for each node for brevity, recall that each node has 26 possible neighbors (3D volume).

Our methodology proceeds in two phases: the training phase and the inference phase. The overview of our methodology is shown in Figure 5.

**3.1 Training Phase** As mentioned earlier, while each CRF can potentially have varying numbers of nodes, the set of parameters ( $\lambda$ ) for all the images is the same. We employed the **UGM package** for learning the parameters of the CRF<sup>3</sup>; this is because it is one of the few packages that can learn a CRF with a large number of parameters, and continuous evidence variables. The number of parameters learned for each CRF is 96, which corresponds to 64 edge features, 16 observation (node) features, and 16 boundary edge features. Thus, a possibly 3 million node CRF can be efficiently represented using 96 parameters; this is one of the main advantages of parameter representation via exponential functions.

We considered three CRF-training approaches:

<sup>3</sup><http://www.di.ens.fr/~mschmidt/Software/UGM.html>

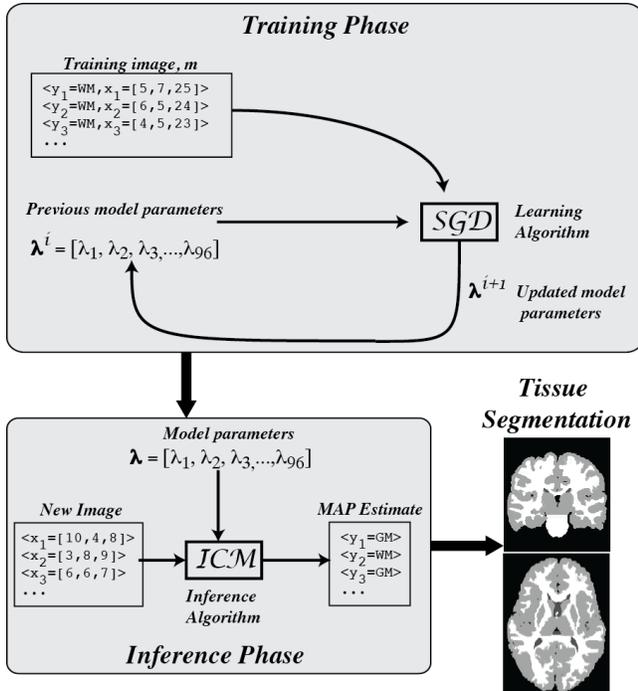


Figure 5: Overview of CRF model training and inference. *SGD* and *ICM* stand for stochastic gradient descent and iterated conditional modes, respectively.

pseudo likelihood [5], L-BFGS [19] and stochastic gradient descent. Because of the size of the graph, the memory requirement for learning using batch methods was prohibitively high. In our experiments, we found that using stochastic gradient descent performed the best compared to other training methods, and consequently employed learning with stochastic gradient descent. Recall that stochastic gradient descent is an on-line algorithm that iterates over each example, computing the gradient with respect to each example. It makes several passes over the training set before converging to the optimal parameters. We used a random ordering of the training images between each iteration. In our experiments, the number of iterations required for convergence was between 200 and 500 iterations. We employed loopy belief propagation (BP) [21, 23] as the inference algorithm for estimating the partition function during training and marginal probabilities *during training*. The only user-tunable parameter is the maximum number of iterations, and this value was set using 5-fold cross validation.

**3.2 Inference Phase** Once the parameters of the CRF are estimated, the next step is to classify the tissue type at each voxel of the image. This problem is posed as obtaining the maximum *a posteriori* (MAP) estimate

over the different voxels i.e.,

$$(3.5) \quad \hat{y} = \arg \max_{y_i} P(y_i = \hat{y}_i | \mathbf{x}_i) \quad \forall i.$$

In order to perform inference, we use iterated conditional models (ICM) [6], which maximizes local conditional probabilities sequentially. The algorithm exploits the notion that neighboring voxels typically are of the same type (GM, WM or CSF) and that each voxel is corrupted with a given probability. Simply put, the aim of ICM is to minimize the within-segment variance by assigning each voxel a specific label, while taking the neighborhood information into account. Thus, a set of neighboring voxels with the same label type will form a “segment” within the image. To avoid reaching local minima, ICM can be used with restarts; we used 30 restarts in our experiments. We preferred ICM over loopy belief propagation for MAP inference because ICM is scalable and fast; additionally, the presence of restarts allowed avoiding local minimums that would otherwise occur due to the use of loopy BP.

## 4 Experimental Setup

Our experiments were designed to answer the following questions:

- Q1: How does the proposed approach compare against atlas-based (knowledge intensive) MRI-image analysis methods?
- Q2: How does the proposed approach perform against state-of-the-art probabilistic methods (atlas-free) for MRI-image analysis?
- Q3: How does the proposed method generalize across different data sets?

Ideally, we would like to compare methods on an Alzheimer’s data set (such as the ADNI study<sup>4</sup>) in the disease prediction task. To our knowledge, there are no publicly available data sets with manual annotations for abnormal brain MRI images.

**4.1 Data Sets** Data was acquired from the Internet Brain Segmentation Repository (IBSR)<sup>5</sup>. IBSR provides two data sets, IBSR V1.0 and IBSR V2.0. IBSR V1.0 consists of 20 low resolution, normal brains. IBSR V2.0 consists of 18 high resolution 1.5mm T1-weighted scans. The scans have been spatially normalized through rotation only, and processed by the Center for Morphometric Analysis (CMA) *AutoSeg* bias field correction routines. Both data sets include manual tissue segmentations by experts, which was used as ground truth.

<sup>4</sup><http://www.adni-info.org/>

<sup>5</sup><http://www.nitrc.org/projects/ibsr>

**4.2 Comparative Algorithms** Other MRI analysis methods we compared against were SPM8 New Segment (SPM8+), VBM8, and FAST. Static Parametric Mapping 8 (SPM8) is a software suite implemented in Matlab for MRI analysis. It includes two tissue segmentation methods, Segment and New Segment. New Segment differs from Segment in that it also classifies non-brain tissues, and performs a post-registration cleanup using a Markov random field (MRF). As New Segment typically outperforms Segment [9], we did not include results from Segment in this paper. For SPM8+, registration into MNI space was first performed using SPM8 coRegister. SPM8+ performs a full bias field correction and tissue segmentation using an atlas-based MAP method. SPM8+ outputs marginal probability maps for each tissue class; these were compiled using the maximum marginal probability given each voxel [3]. SPM8+ was performed using default segmentation parameters, a light bias field correction, and a cleanup MRF of strength 1.

Voxel-Based Morphometry 8 (VBM) uses an atlas-based MAP method combined with partial volume estimation and denoising. VBM was performed using default SPM8 batch parameters. Maximum marginal probabilities for each voxel were used to compile a final segmentation. FAST is a fully automatic tissue segmentation framework within the FSL software suite. FAST version 5.0 was used.

FAST performs segmentation using a hidden Markov Random field fitted through an expectation maximization algorithm while simultaneously performing bias field correction, outputting both tissue probability maps and a compiled final image [30]. The compiled image was used for evaluation.

**4.3 Experiments** A tissue segmentation of the 18 IBSR V2.0 images was performed to evaluate the accuracy of the proposed model. Full leave-one-out cross validation, with a five-fold cross-validation tuning set was used. WM and GM results were compared against FAST, SPM8+, and VBM8. Classification of CSF was also performed, but not included (the reasons for this are explained at the end of this subsection). To demonstrate generalizability of the proposed method, a second test was performed using both IBSR V1.0 and IBSR V2.0. CRFs were trained only using the lower resolution IBSR V1.0 images, and then tested on the higher resolution IBSR V2.0 images.

Segmentation accuracies were evaluated using the Dice coefficient [8]. The Dice coefficient is related to the Jaccard similarity index and  $F_1$ -score in that they are all monotonic with respect to one another. The Dice index is a commonly used measure of segmentation accuracy

in neuroimaging [9]. For two samples  $A$  and  $B$ , the Dice coefficient is defined as

$$(4.6) \quad \frac{A \cup B}{|A| + |B|}.$$

For segmentation analysis this becomes

$$(4.7) \quad \frac{2(TP)}{(TP + FP) + (TP + FN)}$$

where  $TP$  is the true positive rate,  $FP$  is the false positive rate, and  $FN$  is the false negative rate. A Dice coefficient of 1 means that the given segmentation is exactly the same as the ground truth image, while a Dice coefficient of 0 means they are completely non-overlapping.

Finally, we mentioned earlier that the discrimination experiments for CSF were performed but not reported. This is because the IBSR V2.0 manual segmentations do not include certain types of CSF. This missing data causes all algorithms to perform at a Dice index less than 0.4, leading to anomalous results.

## 5 Results and Discussion

In this section, we present the results of the experiments. Figure 6 presents the Dice coefficient for the WM and GM regions respectively. For example, for the WM Dice coefficient, we averaged over all the voxels where the “true” label from the manual segmentation is WM. Higher values would indicate that the WM regions have been predicted more accurately by the model.

**5.1 Comparison to Atlas-Based Methods** As shown in the figure, the proposed approach (denoted as CRF in the graphs) performs significantly better than the atlas-based methods (SPM8+ and VBM) for both WM and GM. Hence, Q1 can be answered affirmatively – the proposed approach is better than the atlas-based methods in isolating the WM and GM regions.

**5.2 Comparison to Atlas-Free Methods** When compared to the state-of-the-art probabilistic atlas-free method (FAST), the CRF method is slightly worse in WM prediction and is slightly better in GM prediction, making its performance comparable to recent approaches for MRI segmentation. Hence, Q2 can be answered neutrally in that the methods are comparable. The key advantage of our method is that *it can be easily implemented on any available (scalable) CRF implementation, and does not require specialized learning and inference modules or hardware*. The FAST method is related to our method as they both use a hidden component for MRFs, but it is well-known that training CRFs is easier than training MRFs. It would be an interesting

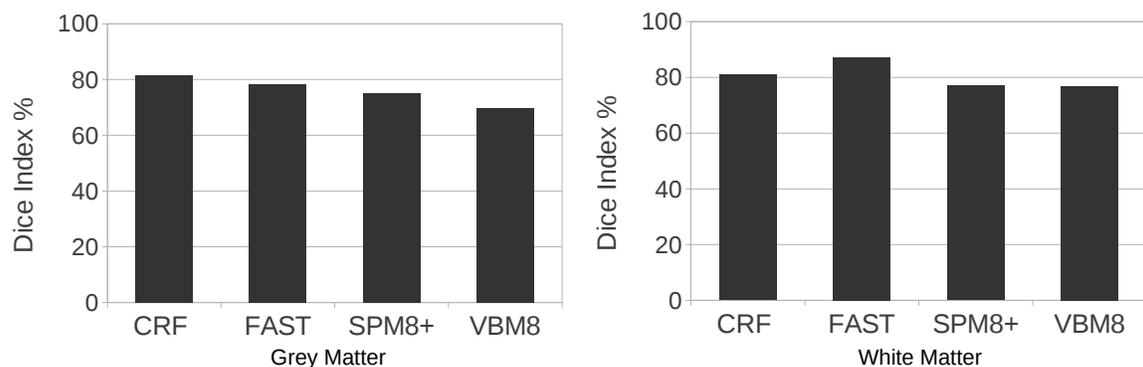


Figure 6: Dice Coefficients for WM and GM predictions after being training and tested on subsets of the IBSR V2.0 data set. These are the results of leave-one-out testing using 5-fold cross validation. The results are averaged over 18 runs of the test set.

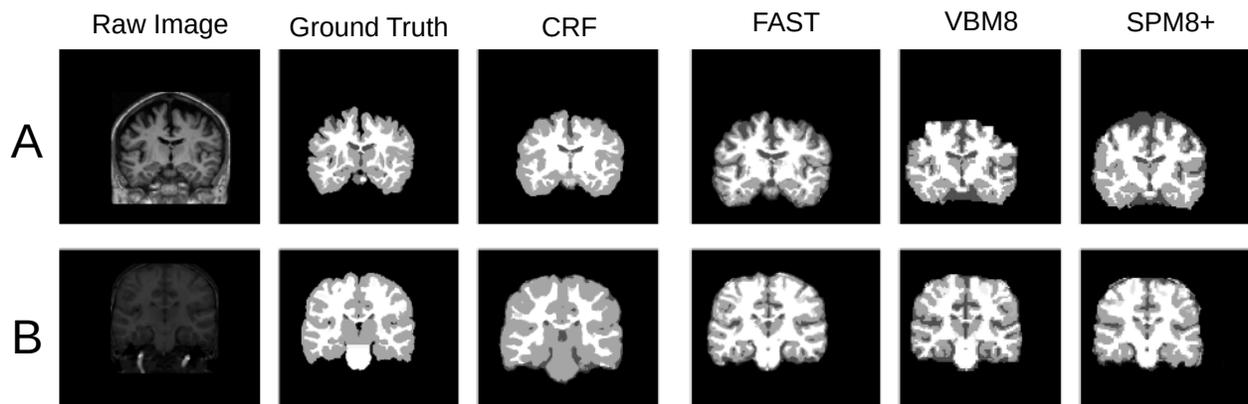


Figure 7: Two example MRI images where our method does very well (A) and poorly (B). In A which contains slight abnormalities, as mentioned in the introduction, our CRF method performs the best along with FAST when compared to atlas-based methods. In B, which is an image with areas of low average intensity due to scanning noise, the CRF method overestimated the grey matter compared to the other methods.

future direction to explore the use of Gaussian mixture models (along the lines of FAST) for CRFs to gain improvements in performance. In addition, EM training did not do well in our method, compared to the stochastic conjugate gradient descent; in fact, it was an order of magnitude slower, while FAST employs EM for training its hidden MRF. Finally, it should be noted that *our method entails very little domain-engineered knowledge* (e.g., bias field correction, expert knowledge constraints, and the use of priors), which FAST and other methods do incorporate. One future direction would be to incorporate these features into the CRF model, which adds increased expressivity to the models. Our initial experiments avoid this as we seek to develop a general im-

age analysis framework that can extend beyond tissue segmentation (e.g., classification of disease; anatomical segmentation).

To understand the performance of our method further, we consider two specific images and present the results in Figure 7. The figure shows two brain images A and B, and the raw image is presented along with the ground truth. As can be seen in A, which has a mildly deformed brain structure, our proposed method and FAST appear to identify the white and gray matter regions correctly. However, the atlas-based methods are very general and are not sensitive to changes in brain structures; they breakdown in such methods. For example, VBM8 predicts some WM

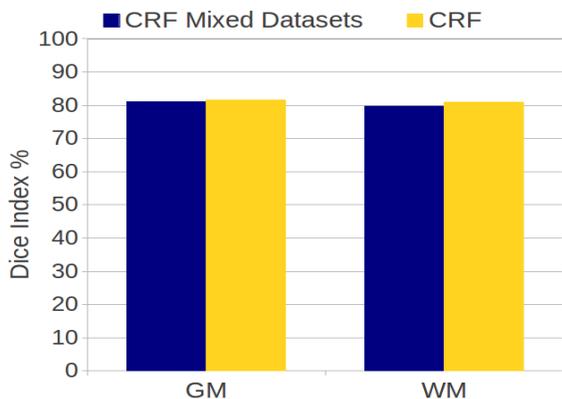


Figure 8: Results of comparing CRF based approach when trained from the same IBSR V2.0 data set (yellow) and when trained on low-resolution IBSR V1.0 data set and evaluated on the high resolution data set (blue). The results are comparable with no statistically significant difference between the scenarios on both the gray matter and white matter predictions. The results are averaged over 5 runs.

regions to be adjacent to the background region (which generally does not happen with most images). When predicting for B, however, because of B having a much lower intensity across the image in this slice when compared to A, our method predicts more GM than actually present in the image. We believe that this is due to the fact that our method models GM very well (as evidenced by the earlier experiment), but when the average intensity is on the lower side compared to the rest of MRI, it causes the method to predict more regions as GM. FAST does not experience this drop in accuracy as it incorporates corrections of these intensity inhomogeneities (generally termed bias field correction) in its framework. Exploring the reason for this mild over fitting remains an interesting future work direction.

**5.3 Comparison Across Data Sets** To answer Q3, we trained the model on low resolution IBSR V1.0 images, and tested them on high resolution IBSR V2.0 images. We compared the results to simply training and testing on the high resolution IBSR V2.0 images. The results are presented in Figure 8, where the former setting is presented in blue and the latter in yellow for both WM and GM prediction. For both tasks, the CRF method generalized quite well across the data sets, given that there is no statistical significance in the difference in performance between both settings. This allows us to answer Q3 affirmatively that the CRF method *can generalize across multiple images quite effectively, even with no co-registration step* between images.

**5.4 Computational Costs** All tests were performed using a 4-core Intel i5-3470S 2.9 GHz processor and 31 GB of RAM. Each fold took approximately 30 minutes to decode new images after training (comparable to other methods). Training took approximately 40 hours to complete. It is important to note that when comparing run times, methods such as SPM or FAST have already been fully trained, and thus it is appropriate to only compare the time it takes the proposed method to decode new images.

## 6 Conclusion and Future Work

As far as we are aware, this is the first work on employing the highly successful framework of CRFs on per voxel based analysis for MRI images, specifically for predicting WM and GM regions in MRI analysis from voxel data. We have demonstrated that we could employ an CRF learner to learn a small number of parameters that are shared by different CRFs. The results were superior to that of atlas-based methods while being comparable to the state-of-the-art MRF based method. When compared to the MRF method, we employ no domain engineered features. We also demonstrated that the resulting classifier allowed for generalization across multiple resolution images.

The most logical next step is to begin to apply our framework to other MRI analysis problem. Anatomical segmentation has posed an especially challenging problem for atlas-free methods, as evidenced the most commonly used MRF based anatomical segmentation method still heavily relies on an atlas [10]. An interesting future direction would be to see if our method can perform atlas-free anatomical segmentation. Another powerful framework for multi-dimensional spatial image analysis is a special type of CRF called the Discriminative Random Field (DRF) method [15]. The key advantage of DRFs is that they allow for domain specific classifiers to model the potential functions. A related work employs support vector machine based classifier for capturing the observation potentials [17] to identify large tumors in MRI images. We on the other hand, consider voxel by voxel data for modeling the brain. It is an interesting direction to explore more expressive classifiers inside our potential functions. Another possible future work is considering and comparing other learning methods [27] for training the CRFs. Finally, using the results of the image analysis for direct prediction of events such as onset of Alzheimer’s is an interesting and exciting future research possibility. We believe that the real impact of anatomical segmentation can be realized by combining their output with powerful classifiers.

## References

- [1] P. Aljabar, R. Heckemann, A. Hammers, J. Hajnal, and D. Rueckert. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *NeuroImage*, 46(3):726–738, 2009.

- [2] J. Ashburner and K. J. Friston. Voxel-based morphometry - the methods. *NeuroImage*, 11(6):805–821, 2000.
- [3] J. Ashburner and K. J. Friston. Unified segmentation. *NeuroImage*, 26(3):839–851, 2005.
- [4] M. Balafar, A. Ramli, M. Saripan, and S. Mashohor. Review of brain MRI image segmentation methods. *Artificial Intelligence Review*, 33(3):261–274, 2010.
- [5] J. Besag. Statistical Analysis of Non-Lattice Data. *The Statistician*, 24(3):179–195, 1975.
- [6] J. Besag. On the Statistical Analysis of Dirty Pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(3):259–302, 1986.
- [7] P.-L. Chang and W.-G. Teng. Exploiting the self-organizing map for medical image segmentation. In *Computer-Based Medical Systems, Twentieth IEEE Intl. Symp. on, CBMS '07*, pages 281–288, 2007.
- [8] L. R. Dice. Measures of the Amount of Ecologic Association Between Species. *Ecology*, 26(3):297–302, 1945.
- [9] L. D. Eggert, J. Sommer, A. Jansen, T. Kircher, and C. Konrad. Accuracy and reliability of automated gray matter segmentation pathways on real and simulated structural magnetic resonance images of the human brain. *PLoS ONE*, 7(9), 09 2012.
- [10] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. M. Dale. Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuron*, 33(3):341–355, 2002.
- [11] N. Friedman and S. Russell. Image segmentation in video sequences: a probabilistic approach. In *Proc. Thirteenth Conf. on Uncertainty in Artificial Intelligence, UAI'97*, pages 175–181, 1997.
- [12] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Trans.*, PAMI-6(6):721–741, 1984.
- [13] X. He, R. S. Zemel, and M. A. Carreira-Perpiñán. Multiscale conditional random fields for image labeling. In *Proc. 2004 IEEE Conf. on Computer Vision and Pattern Recognition, CVPR'04*, pages 695–703, 2004.
- [14] K. Held, E. Kops, B. Krause, W. Wells, R. Kikinis, and H. Muller-Gartner. Markov random field segmentation of brain MR images. *Medical Imaging, IEEE Trans.*, 16(6):878–886, 1997.
- [15] S. Kumar and M. Hebert. Discriminative fields for modeling spatial dependencies in natural images. In *Advances in Neural Information Processing Systems 16, NIPS '03*, 2003.
- [16] J. D. Lafferty, A. McCallum, and F. C. N. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proc. Eighteenth Intl. Conf. on Machine Learning, ICML '01*, pages 282–289, 2001.
- [17] C.-H. Lee, M. Schmidt, A. Murtha, A. Bistriz, J. Sander, and R. Greiner. Segmenting brain tumors with conditional random fields and support vector machines. In Y. Liu, T. Jiang, and C. Zhang, editors, *Computer Vision for Biomedical Image Applications*, volume 3765 of *LNCIS*, pages 469–478. 2005.
- [18] S. Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer, 3rd edition, 2009.
- [19] D. C. Liu and J. Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical Programming*, 45:503–528, 1989.
- [20] J. A. Maldjian, P. J. Laurienti, R. A. Kraft, and J. H. Burdette. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19(3):1233–1239, 2003.
- [21] K. P. Murphy, Y. Weiss, and M. I. Jordan. Loopy belief propagation for approximate inference: an empirical study. In *Proc. Fifteenth Conf. on Uncertainty in Artificial Intelligence, UAI '99*, pages 467–475, 1999.
- [22] S. Natarajan, B. Saha, S. Joshi, A. Edwards, T. Khot, E. M. Davenport, K. Kersting, C. T. Whitlow, and J. A. Maldjian. Relational learning helps in three-way classification of Alzheimer patients from structural magnetic resonance images of the brain. *Intl. Journal of Machine Learning and Cybernetics*, pages 1–11, 2013.
- [23] J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.
- [24] A. Quattoni, M. Collins, and T. Darrell. Conditional random fields for object recognition. In *Advances in Neural Information Processing Systems 17, NIPS '04*, pages 1097–1104, 2004.
- [25] T. Song, M. Jamshidi, R. Lee, and M. Huang. A modified probabilistic neural network for partial volume segmentation in brain MR image. *Neural Networks, IEEE Trans.*, 18(5):1424–1432, 2007.
- [26] L. Sun, R. Patel, J. Liu, K. Chen, T. Wu, J. Li, E. Reiman, and J. Ye. Mining brain region connectivity for Alzheimer's disease study via sparse inverse covariance estimation. In *Proc. 15th Intl. Conf. on Knowledge Discovery and Data Mining, KDD '09*, pages 1335–1344, 2009.
- [27] C. Sutton and A. McCallum. An introduction to conditional random fields, 2010. arxiv:1011.4088.
- [28] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classification of MR images of the brain. *Medical Imaging, IEEE Trans.*, 18(10):897–908, 1999.
- [29] J. Ye, K. Chen, T. Wu, J. Li, Z. Zhao, R. Patel, M. Bae, R. Janardan, H. Liu, G. Alexander, and E. Reiman. Heterogeneous data fusion for Alzheimer's disease study. In *Proc. 14th Intl. Conf. on Knowledge Discovery and Data Mining, KDD '08*, pages 1025–1033, 2008.
- [30] Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *Medical Imaging, IEEE Trans.*, 20(1):45–57, 2001.