Conditional Random Fields for brain tissue segmentation

Chris S. Magnano[1], Ameet Soni[1], Sriraam Natarajan[2], and Gautam Kunapuli[3]

[1] Swarthmore College, Computer Science, {cmagnan1,soni}@cs.swarthmore.edu
[2] Indiana University, School of Informatics and Computing, natarasr@indiana.edu
[3] UtopiaCompression Corporation, gautam@utopiacompression.com

Abstract. Current atlas-based methods for MRI analysis assume brain images map to a "normal" template. This assumption, however, does not hold when analyzing abnormal brain shapes or disease states. We propose a discriminative-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. Our algorithm generalizes well across data sets and is less susceptible to outliers, while relying on minimal prior knowledge relative to atlas-based techniques. These results provide a promising framework for future application on disease classification and atlas-free anatomical segmentation.

1 Introduction

This paper presents a novel framework for extracting useful information from structural MRI images. Many image techniques such as thresholding, region growing, and clustering have been previously used for medical imaging. As discussed in Balafar et al. [4], thresholding methods are not ideal since the distribution of intensities in medical images is quite complex. Atlas-based segmentation methods, on the other hand, parcellate the MRI data into different anatomical regions using a standard brain template. Effective MRI-based diagnosis 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; all brains are segmented into the same set of regions and assume normal characteristics. Our work aims to find an atlas-free, discriminative machine-learning-based approach to segment brain images and detect abnormalities that could *aid in classification of diseases from images.* Recent work [17] on predicting the incidence of AD using atlas-free segmentation combined with a classifier demonstrates that such an approach can produce superior diagnostic performance.

While several diverse paradigms exist for image segmentation, we focus on probabilistic models, as they have been used successfully in many image segmentation tasks and provide a mechanism for handling noise and complex structures in data[6, 22]. Markov Random Fields (MRFs) [10], for example, have been applied to a wide variety of tasks including texture analysis and image restoration,

among others [15] as well as brain MRI segmentation [12, 25] and tissue classification [24]. Recent years have seen the emergence of Conditional Random Fields (CRFs) [14, 23], 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 [11] and object recognition [19]. The work of Kumar and Hebert [13] 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. We propose the first of its kind, fully-CRF-based framework for structural-MRI-image analysis, and apply this to the task of volumetric segmentation for 3-dimensional data. We apply our approach to standard brain repository data sets and show that our method achieves superior performance to popular atlas-based methods and comparable performance to the state-of-the-art methods.

2   CRFs for MRI Image analysis

We now outline our proposed method for identifying relevant regions from MRI images. A thorough introduction to CRFs is provided by Sutton and McCallum [23]. Here, we seek to model our MRI task using the following model:

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

where $Z(\mathbf{x})$ is a normalization factor, $\mathbf{y}$ is the set of hidden variables (one for each voxel's tissue type; i.e., $y_i = \{\texttt{WM}, \texttt{GM}, \texttt{CSF}\}$ for gray matter, white matter and cerebral spinal fluid), $\mathbf{x}$ is the set of observed variables for each voxel (i.e., voxel intensity, neighborhood average intensity, and distance to center), f_k defines the kth feature function, and λ_k is the corresponding weight given to that feature. We chose to model each voxel as being connected to 26 neighboring voxels (i.e., y_i is connected to each y_j in $3 \times 3 \times 3$ neighborhood of hidden variables).

The input to the algorithm is a CRF with defined structure, a set of training examples in the form of MRI images (observations $\mathbf{x}$) with corresponding tissue segmentation (ground truth labels $\mathbf{y}$). Each MRI image has about 3 million voxels yielding a corresponding CRF with 3 million hidden variables.

Estimating the conditional distribution corresponds to estimating λ for the feature set. While the CRF structure for each image could be different due to different brain sizes, they all share the same set of parameters due to the features being the same for each node in the CRF. Our methodology proceeds in two phases: the training phase and the inference phase and is shown in Figure 2.

Training Phase: As mentioned earlier, while each CRF can potentially have varying numbers of nodes, the set of parameters (λ) for all the images is the same. We employed the `UGM package` [21] for learning the parameters of the CRF, 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 features, and 16 boundary edge features. Thus, a possibly 3 million node CRF can be efficiently represented using 96 parameters.
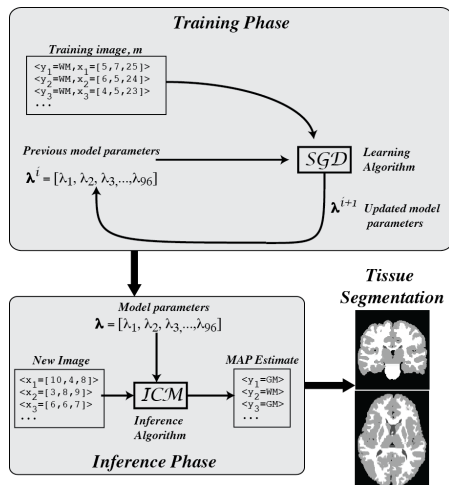
Training Phase

Training image, m

```
<y1=WM,x1=[5,7,25]>
<y2=WM,x2=[6,5,24]>
<y3=WM,x3=[4,5,23]>
...
```

Previous model parameters

$\boldsymbol{\lambda}^i = [\lambda_1, \lambda_2, \lambda_3, ..., \lambda_{96}]$

SGD *Learning Algorithm*

$\boldsymbol{\lambda}^{i+1}$ *Updated model parameters*

Tissue Segmentation

Model parameters

$\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \lambda_3, ..., \lambda_{96}]$

New Image

```
<x1=[10,4,8]>
<x2=[3,8,9]>
<x3=[6,6,7]>
...
```

ICM *Inference Algorithm*

MAP Estimate

```
<y1=GM>
<y2=WM>
<y3=GM>
...
```

Inference Phase

Fig. 1. Overview of CRF model training and inference. SGD and ICM stand for stochastic gradient descent and iterated conditional modes, respectively.

We considered three primary CRF-training approaches: pseudo likelihood, L-BFGS 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. Stochastic gradient descent is an online 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 use a random ordering of the training images with the optimal number of iterations ranging from 200 and 500 iterations. We employed loopy belief propagation (BP) [16, 18] as the inference algorithm for estimating the partition function during training and marginal probabilities The only user-tunable parameter is the maximum number of iterations, and this value was set using 5-fold cross validation.

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., $\hat{\mathbf{y}} = \arg\max_{y_i} \quad P(y_i = \hat{y}_i \,|\, \mathbf{x}_i) \quad \forall i$. In order to perform inference, we use iterated conditional models (ICM) [5], 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.

3 Experimental Setup

We aim to answer the following questions: **Q1:** How does the proposed approach compare against the atlas-based (knowledge intensive) MRI image analysis methods? **Q2:** How does the CRF method perform against the state-of-

the-art probabilistic method (atlas-free) for MRI analysis? **Q3:** How does the proposed method generalize across different data sets? Ideally, we would like to compare the methods on an Alzheimer's data set (such as the one from ADNI) in the disease prediction task, but, as far as we are aware, there are no publicly available data sets with manual annotations for abnormal brain MRI images.

Dataset: Data was acquired from the Internet Brain Segmentation Repository(IBSR) [20]. 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' biasfield correction routines. Both data sets include manual tissue segmentation by experts which were used as a ground truth.

Other Algorithms: Results were compared against Voxel-Based Morphometry (VBM8) [2] , SPM8 New Segment (SPM8+) [3, 8], and FAST [25]. SPM8+ was performed using default segmentation parameters, a light bias field correction, and a cleanup MRF of strength 1. VBM uses an atlas based maximum posterior probability method combined with partial volume estimation and denoising. VBM was performed using default SPM8 batch parameters. FAST is a tissue segmentation software within the FSL software suite that uses a hidden Markov random field fitted through an expectation maximization algorithm [25].

Experiments: A tissue segmentation of the 18 IBSR V2.0 images was performed to evaluate the accuracy of our CRF framework. Full leave one out cross validation with a five fold cross validation tuning set was used for the CRF framework. Results were compared against FAST, SPM8+, and VBM8. To demonstrate the robustness of the CRF framework, a second test was performed where the CRF was trained only using the lower resolution IBSR V1.0 images and then deployed on the higher resolution IBSR V2.0 images. Segmentation accuracies were evaluated using the Dice coefficient [7]. 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 measures the degree of spatial overlap between two sets of voxels, and varies from 0 (no overlap) to 1 (complete overlap). The Dice index is a commonly used measure of segmentation accuracy in neuroimaging [8].

4   Results and Discussion

Figure 2 present the Dice coefficient for the WM regions 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. Hence, higher values would indicate that the WM regions have been predicted more accurately by the model. As can be seen from the figure, both for WM and GM, the proposed approach (denoted as CRF in the graphs) performs significantly better than the atlas-based methods (SPM8 and VBM). Hence, Q1 can be answered affirmatively that the proposed approach is better than the atlas-based methods in isolating the WM and GM regions.
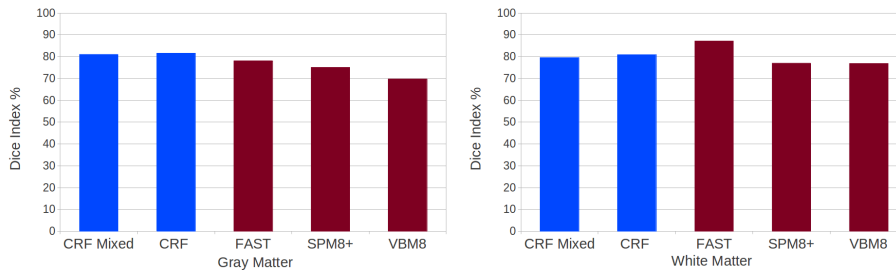
Fig. 2. Dice Coefficients for WM and GM predictions averaged over the 18 images in the `IBSR V2.0` data set. Our method is presented in blue while other methods are presented in red. **CRF** is our proposed model, trained on the `IBSR V2.0` data while **CRF mixed** is the result of training on the low resolution `IBSR V1.0` data.

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 future direction to explore the use of Gaussian mixture models (along the lines of FAST) for CRFs to gain improvements in performance. In addition, 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 image 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 3. 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 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
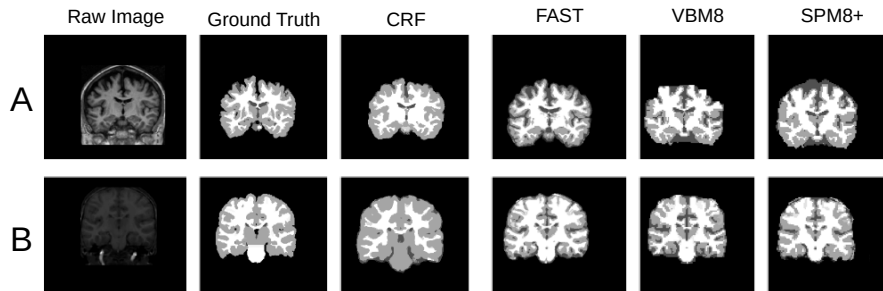
| | Raw Image | Ground Truth | CRF | FAST | VBM8 | SPM8+ |

Fig. 3. Two example MRI images where our method does very well (A) and poorly (B). In A, which contains slight abnormalities, our CRF method performs the best along with FAST when compared to atlas-based methods. In B, an image with areas of low average intensity due to scanning noise, the CRF method overestimated the gray matter compared to the other methods. The low contrast in B gives an advantage to methods with integrated domain knowledge.

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.

To answer Q3, we trained the model on low resolution `IBSR V1.0` images, and tested them on high resolution `IBSR V2.0` images. Figure 2 shows the result as **CRF Mixed**. Importantly, the results for this experiment are nearly identical to the normal set up (leave one out training/testing on `IBSR V2.0` alone). This reinforces the notion that a trained CRF method generalizes quite well across the data sets despite different resolutions. This allows us to answer Q3 affirmatively, an important result when considering an extension to disease-specific data sets.

5 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. These results provide an framework for future directions of work, which aims to identify intermediate disease states with higher accuracy (e.g., mild cognitive impairment) [17] and improve upon anatomical segmentation techniques which rely heavily upon brain atlases [9].

References

1. P. Aljabar, R.A. Heckemann, A. Hammers, J.V. 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. John Ashburner and Karl J. Friston. Voxel-based morphometry – the methods. *NeuroImage*, 11(6):805–821, 2000.

3. John Ashburner and Karl J. Friston. Unified segmentation. *NeuroImage*, 26(3):839–851, 2005.

4. M.A. Balafar, A.R. Ramli, M.I. Saripan, and S. Mashohor. Review of brain MRI image segmentation methods. *Artificial Intelligence Review*, 33(3):261–274, 2010.

5. Julian Besag. On the Statistical Analysis of Dirty Pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(3):259–302, 1986.

6. Ping-Lin Chang and Wei-Guang 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.

7. L. R. Dice. Measures of the Amount of Ecologic Association Between Species. *Ecology*, 26(3):297–302, 1945.

8. Lucas D. Eggert, Jens Sommer, Andreas Jansen, Tilo Kircher, and Carsten 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.

9. Bruce Fischl, David H. Salat, Evelina Busa, Marilyn Albert, Megan Dieterich, Christian Haselgrove, Andre van der Kouwe, Ron Killiany, David Kennedy, Shuna Klaveness, Albert Montillo, Nikos Makris, Bruce Rosen, and Anders M. Dale. Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuron*, 33(3):341–355, 2002.

10. Stuart 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.

11. Xuming He, Richard S. Zemel, and Miguel Á. 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.

12. K. Held, E.R. Kops, B.J. Krause, W.M. 3rd Wells, R. Kikinis, and H.W. Muller-Gartner. Markov random field segmentation of brain MR images. *Medical Imaging, IEEE Trans.*, 16(6):878–886, 1997.

13. Sanjiv Kumar and Martial Hebert. Discriminative fields for modeling spatial dependencies in natural images. In *Advances in Neural Information Processing Systems 16*, NIPS '03, 2003.

14. John D. Lafferty, Andrew McCallum, and Fernando 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.

15. Stan Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer, 3rd edition, 2009.

16. Kevin P. Murphy, Yair Weiss, and Michael 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.

17. Sriraam Natarajan, Baidya Saha, Saket Joshi, Adam Edwards, Tushar Khot, Elizabeth M. Davenport, Kristian Kersting, Christopher T. Whitlow, and Joseph 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.

18. Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.

19. Ariadna Quattoni, Michael Collins, and Trevor Darrell. Conditional random fields for object recognition. In *Advances in Neural Information Processing Systems 17*, NIPS '04, pages 1097–1104, 2004.

20. Internet Brain Segmentation Repository. http://www.cma.mgh.harvard.edu/ibsr/.

21. Mark Schmidt. http://www.di.ens.fr/ mschmidt/Software/UGM.html.

22. Tao Song, M.M. Jamshidi, R.R. Lee, and Mingxiong Huang. A modified probabilistic neural network for partial volume segmentation in brain MR image. *Neural Networks, IEEE Trans.*, 18(5):1424–1432, 2007.

23. Charles Sutton and Andrew McCallum. An introduction to conditional random fields. http://arxiv.org/abs/1011.4088, 2010. arxiv:1011.4088.

24. 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.

25. Yongyue 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.