Chapter 5
Machine Learning for Brain Image Segmentation

Jonathan Morra
University of California Los Angeles, USA

Zhuowen Tu
University of California Los Angeles, USA

Arthur Toga
University of California Los Angeles, USA

Paul Thompson
University of California Los Angeles, USA

ABSTRACT
In this chapter, the authors review a variety of algorithms developed by different groups for automatically segmenting structures in medical images, such as brain MRI scans. Some of the simpler methods, based on active contours, deformable image registration, and anisotropic Markov random fields, have known weaknesses, which can be largely overcome by learning methods that better encode knowledge on anatomical variability. The authors show how the anatomical segmentation problem may be re-cast in a Bayesian framework. They then present several different learning techniques increasing in complexity until they derive two algorithms recently proposed by the authors. The authors show how these automated algorithms are validated empirically, by comparison with segmentations by experts, which serve as independent ground truth, and in terms of their power to detect disease effects in Alzheimer’s disease. They show how these methods can be used to investigate factors that influence disease progression in databases of thousands of images. Finally the authors indicate some promising directions for future work.

INTRODUCTION
Automated analysis of brain scans is increasingly important as the cost of acquiring a brain scan decreases, and the frequency of their use increases. Drug trials and genetic studies often collect hundreds
or thousands of images, and efficient algorithms are increasingly needed to compute morphometric statistics. Automated segmentation has been successfully applied to magnetic resonance images (MRI), which are used clinically to examine disease effects. Research studies of large-scale image databases now survey thousands of images at once. These population-based image analyses have discovered how diseases spread in the living brain over time (P. M. Thompson et al., 2003), which medications best resist brain changes in disease (Jack, Petersen et al., 2008; P. M. Thompson et al., 2008), and have discovered specific genes that protect the brain from illness (Hua et al., 2008), or increase the risk for disease (Leow et al., 2008; J. Morra, Z. Tu, L. G. Apostolova, A. Green et al., 2008b; J. Morra, Z. Tu, L. G. Apostolova, A. Green, C. Avedissian, S. Madsen, N. Parikshak, X. Hua, A. Toga, C. Jack, N. Schuff, M. W. Weiner et al., 2008). All of these studies have been accelerated by learning approaches that identify and analyze features in brain images automatically (Fischl et al., 2002; Grenander & Miller, 1998).

MRI scans can be automatically analyzed using a sequence of several steps, including intensity normalization, registration to a common template, segmentation of specific substructures, and statistical analysis. In this chapter, we will focus on current trends for segmenting brain structures on MRI, focusing specifically on learning methods. Most of these approaches are somewhat generic, and have been used to segment images of the heart, liver, lungs, and other organs. They are also applicable in principle to other types of biomedical images, such as computed tomography or histology (Pitiot, Delingette, & Thompson, 2005).

In MRI studies, automated segmentations have been used to compute volumetric measures or shape statistics for specific brain regions, in studies of Alzheimer’s Disease (Apostolova et al., 2007; Clare, Woods, Moniz Cook, Orrell, & Spector, 2003; Csernansky et al., 2000; J. Morra, Z. Tu, L. G. Apostolova, A. Green et al., 2008a, 2008b; J. Morra, Z. Tu, L. G. Apostolova, A. Green, C. Avedissian, S. Madsen, N. Parikshak, X. Hua, A. Toga, C. Jack, N. Schuff, M. W. Weiner et al., 2008), epilepsy (Lin et al., 2005), childhood development (Gogtay et al., 2006), autism (Nicolson et al., 2006), drug-related degeneration in methamphetamine users (Thompson, Hayashi, Simon et al., 2004), and effects of lithium treatment in bipolar illness (Bearden et al., 2007). Figure 1 shows an example of a subject’s brain MRI segmented both by hand and automatically.

Anatomical segmentation is a key step in many of these imaging projects, but most studies still rely on manual segmentations by experts, who delineate each region of interest (ROI) in consecutive sections of each subject’s 3D MRI scan (Apostolova et al., 2006; Schuff et al., 2008). This is time consuming, especially in very large studies. For instance, the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (Jack, Bernstein et al., 2008) is a longitudinal study of 800 subjects scanned five times. Assuming it takes about 2 hours to manually segment the hippocampus from an MRI, then segmenting the hippocampus for all subjects in ADNI would take 2 hours x 2 hippocampi per individual x 800 subjects x 5 time points = 16,000 man-hours for just the hippocampus in this study; clearly this process needs to be automated.

The goal of this chapter is to give an overview of the general principles of image segmentation based on learning. We introduce various methods, increasing in complexity, finally describing a state-of-the-art segmentation algorithm that overcomes several limitations of prior methods. Throughout, we discuss validations that evaluate the accuracy and reproducibility of the segmentations. Finally, we highlight some directions for future research.