Optimal Atlas Selection Methods for Automatic Segmentation of Infant Brains

Alexander Guo, Lilla Zollei

INTRODUCTION. Although the premise of automatic brain segmentation in adult brains has been researched extensively, similar automatic segmentation in infant brains has received relatively little attention. Such automatic segmentations in infant brains, which have greater variability, if performed well, will allow researchers and medical professionals to increase segmentation speed and efficiency while maintaining accuracy, allowing less labor to be devoted on tasks that can be handled automatically. Automatic newborn brain segmentation will additionally lead to greater advances in the study of infant brain development. This study investigates the benefits of utilizing multi-atlas segmentation (MAS) to further ensure segmentation accuracy.

METHOD. A multi-atlas selection framework to automatically segment newborn brains with Magnetic Resonance Imaging (MRI) images was proposed. Fourteen subjects were utilized, complete with their structural brain MRI images and the corresponding manual labels, of which nine are from Boston Children's Hospital and five are from South Africa. Of the 14 subjects, seven different atlas selection methods were proposed and implemented to perform training as atlas selection is a crucial step in segmentation and it is thus necessary to test a variety of methods to obtain the best results. A comparison of results using the Dice coefficient for all seven methods was conducted. These seven methods are as follows. 1) Using all thirteen other atlases for training. 2) Using all nine BCH atlases for training (eight if test subject is BCH). 3) Using all five South African atlases for training (four if test subject is South African). 4) Atlas selection by computing the similarity between each pair of the thirteen training atlases through an average of two similarity features (age and mutual information), and using the pairwise similarity with other subjects for each subject as features for k-means clustering. Similarity between the test subject and all other subjects were also computed, and all atlases in the closest cluster to the test subject were selected for training. 5) Similar to the previous method, with the difference being that three similarity features were used, including age, mutual information, and volume, with a different weighting scheme. 6) Similar to the previous method, with the difference being that instead of selecting all atlases in the closest cluster, the closest atlas in each cluster was selected. 7) Similar to the previous method, with the difference being that instead of selecting the closest atlas in each cluster, the five closest atlases to the test subject were selected regardless of clustering. In addition, number of closest subjects from method 7) were varied between 1 and 13 to determine the optimal number of closest subjects.

RESULTS. A weighted average of the Dice coefficients of all labels was computed for each of the atlas selection methods, based on volume of the respective brain region, in addition to a simple unweighted average of the individual Dice coefficients for each label. These results show that, on average, using all subjects for training was optimal in both label averages in addition to 22 of the 30 individual labels, with Dice coefficients of 80.0% and 70.9% for the weighted and unweighted averages, respectively. This is compared to the averages across all subjects and selection methods of 80.3% and 72.1%, respectively. However, these values from using all data subjects for training were only 2.2% and 0.8% higher than the average Dice coefficients for the last-ranked training set group, respectively (method 2 and method 7, respectively). Additionally,

it was found that most labels and subjects displayed an increasing trend of Dice values through an increase in number of closest subjects from atlas selection (method 7).

CONCLUSION. The results show that although the use of all subjects for training may be optimal, the results are not statistically significant enough to definitively justify the use of one training set over another. Thus, this study has provided additional guidance for the optimal method of atlas selection for automatic segmentation of newborn brains and has demonstrated the feasibility of a MAS atlas selection framework.