Spring 2016

An automated tissue classification pipeline for magnetic resonance images of infant brains using age-specific atlases and level set segmentation

Andrew Metzger
University of Iowa

AN AUTOMATED TISSUE CLASSIFICATION PIPELINE FOR MAGNETIC RESONANCE IMAGES OF INFANT BRAINS USING AGE-SPECIFIC ATLASES AND LEVEL SET SEGMENTATION

by

Andrew Metzger

A thesis submitted in partial fulfillment
of the requirements for the Master of Science
degree in Biomedical Engineering in the
Graduate College of
The University of Iowa

May 2016

Thesis Supervisor: Associate Professor Vincent Magnotta

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

MASTER'S THESIS

_____

This is to certify that the Master's thesis of

Andrew Metzger

has been approved by the Examining Committee for
the thesis requirement for the Master of Science degree
in Biomedical Engineering at the May 2016 graduation.

Thesis Committee: _____
Vincent Magnotta, Thesis Supervisor

_____
Hans Johnson

_____
Mathews Jacob

_____
Dan Thedens

_____
Edwin Dove

_____
Joseph Reinhardt

~ To Marla, Autumn, & Kori ~

ACKNOWLEDGEMENTS

I have been fortunate to have the support of many talented individuals. First and foremost, to Drs. Peg Nopoulos and Vincent Magnotta, thank you for affording me the invaluable opportunity of working with you and your team. To Amanda and Sasha, it has been a great pleasure to share in this exciting project with you. To Dr. Mathews Jacob, the knowledge I have gained under your instruction has been the inspiration of much of my current work and future goals. To Dr. Dan Thedens, your vast expertise and generosity of time is appreciated more than you know. To my lab mates, Cam, Nana, Michael, Merry, and Casey, thank you for the brainstorming, troubleshooting, and camaraderie. Finally, to the members of the Scalable Informatics for Neuroscience, Processing, and Software Engineering lab, Dr. Hans Johnson, Regina, Ali, David, and Jessica, the chance to work with you and your tools has made possible and greatly accelerated any success this work can claim. Funding was provided by the National Institute for Heart, Lung and Blood (NHLBI) Grant 2P01 HL046925 “Preterm Transfusions: Brain Structure and Function Outcomes.”

ABSTRACT

Quantifying tissue volumes in pediatric brains from magnetic resonance (MR) images can provide insight into etiology and onset of neurological disease. Unbiased volumetric analysis can be applied to population studies using automated image processing. Standard segmentation strategies using adult atlases fail to account for varying tissue contrasts and types associated with the rapid growth and maturational changes seen in early neurodevelopment. The goal of this project was to develop an automated pipeline and two age-specific atlases capable of providing accurate tissue classification to address these challenges.

The automated pipeline consisted of a stepwise initial atlas-to-subject registration, expectation maximization (EM) atlas based segmentation, and a post-processing level set segmentation for improved white/gray matter separation. This level set segmentation is a 3D and multiphase adaptation of a 2D method intended for use on images with the types of intensity inhomogeneities found in MR images.

The initial tissue maps required to determine spatial priors for the one-year-old atlas were created by manually cleaning the results of an adult atlas and the automated pipeline. Additional tissue maps were incrementally added until the spatial priors were sufficiently representative. The neonate atlas was similarly created, starting with the one-year-old atlas.

PUBLIC ABSTRACT

Automated image processing can provide unbiased data in large scale population studies to contribute insight into the progression and onset of neurologic disease. Due to the rapid growth and development of the human brain in the first few years of life, automated medical image processing of magnetic resonance (MR) images from infant subjects to investigate neurological conditions is especially challenging. The goal of this work was to create methods and tools suitable to overcome these challenges.

An atlas is a collection of prior knowledge that is used in tissue classification. Because of the fast growth of the brain, a collection of knowledge that is representative of a limited age group can be more precise and useful. For this reason, an atlas was created for the use of two age groups: neonate (images taken shortly after birth) and one-year-old subjects (images taken at 10-18 months of age). To determine the likelihood of a tissue based on its relative spatial positioning and expected image intensity, a set of subjects from each age group were segmented and manually corrected.

MR images of these age groups also suffer from excessive noise and motion artifacts. This can complicate the differentiation of white and gray matter in the cortex. To address this issue, a level set segmentation method, which can be particularly robust to the complex topologies found in the cortex, was adapted to correct tissue classification in an additional post-processing step.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Motivation

The human brain undergoes rapid growth and development within the first few years following parturition.[1] Congenital and environmental conditions can delay or disrupt normal neural development during this time. In a healthy subject, intracranial structures increase in volume and mature at identified tissue specific rates.[2] Quantifying tissue volumes at this stage in development can provide important insight into numerous neurological diseases affecting brain development in pediatric populations. By detecting structural irregularities amongst populations of interest, unrealized avenues of research can be identified, onset of disease can be determined more accurately, and treatment can be more appropriately timed.

Volumetric analysis applied to large population studies are used to accurately detect small differences between groups of interest. When an automated pipeline is available, these studies are considerably more practical and benefit from inherently unbiased results. Results can be regionally isolated to determine functionally significant findings.[3] In addition, automated tissue classification provides convenient groundwork for further investigational methods such as surface analysis and fiber tracking.

1.2 Background

Automated tissue classification of infant brains pose several unique challenges. Standard imaging protocols and systems are designed for adult brains. While pulse sequences continue to be optimized for infants[4,5], this often translates to a considerable loss of signal to noise ratio (SNR) when imaging a smaller subject. Scanning must be

done while the subject is asleep, and artifacts due to motion are characteristically common leading to a higher proportion of unusable images. Figure 1 shows an axial slice from typical T1 weighted image of an adult, one-year-old, and neonatal subjects.

Figure 1: Axial slices of a T1 weighted image of an adult (A), one-year-old (B), and neonatal (C) subject.

As the shape, size and magnetic resonance (MR) properties of anatomical structures are changing rapidly, any prior spatial or intensity knowledge must be age-specific to accurately represent tissue likelihoods. Cortical asymmetries and sexual dimorphisms which are more pronounced at this age increase variability as well.[6] Furthermore, gradual myelination leads to a reversal in T1 and T2 contrasts between cortical white matter and the adjacent gray matter within the first year of development[7] which can lead to ambiguous or poorly defined tissue boundaries. This begets the need for an additional tissue type "unmyelinated white matter" due to the drastic change in contrast.

1.2.1 Modern Approaches in Tissue Classification.

Determining volumes and region boundaries from medical images has been a prominent problem in the fields of computer vision and image processing. As such many strategies have been extensively developed to classify tissue from MR images of the brain.

Thresholding, one of the most fundamental approaches to image segmentation, assigns binary labels based on the intensity at each voxel and a set intensity inequality. Choosing an appropriate threshold can be automated with statistical approaches such as Otsu's method[8] and linear discriminant analysis.[9] This approach has been used in early work to classify broad tissue types from magnetic resonance images for 3D modeling of the brain.[10] Thresholding it often a useful initial or intermediate step, but is limited in use because it does not classically consider spatial information in determining regions. This is especially important when dealing with MR images where intensities can vary due to local bias fields, as well different tissue types of interest that exhibit similar average intensities.

Artificial Neural Networks (ANN) provide one alternative to tissue classification using thresholding based on classical statistics. ANN's resemble biological learning by modeling large networks of parallelized processors or nodes.[11] Weighting of the nodes is adjusted automatically while processing training sets which prime the ANN for segmentation of similar images. These methods are especially adept at reducing dimensionality of images in a sophisticated manner, which is ideal for tissue classification.[12] For example, there exist several ANN methods of image segmentation by recognizing patterns of texture within a region.[12,13] Neural networks excel in automating

or semi-automating much of the analysis that would typically be applied by experts readers of MR images.[14] These advantages will continue to affect computational costs as parallel processing capacity improves. These methods are often limited by the quality and quantity of a priori knowledge that exist in training sets.

Atlas-based segmentation aims to apply spatial and intensity likelihoods derived from a collection of cleaned tissue maps to classify voxels in an image. The ideal atlas completely represents all structural variability within a given population. To approximate this as accurately as possible, the classification maps of a representative group of images are combined to create spatial probabilities. In order to apply those spatial probabilities to a subject of interest, a common coordinate space between the atlas and the image of interest must be determined through image registration. This strategy has been used extensively for tissue classification in the human brain.[19] A derivative method, multi-atlas segmentation, attempts to further improve the accuracy of the application of spatial probability by individually registering each prior classified images to the image of interest before combining classification maps to determine spatial probabilities.[19]

Clustering methods attempt to classify images by iterating between determining the properties each class exhibits as well as the classification of each image element. Expectation maximization (EM) algorithms are a class of clustering methods that assume the intensity profile of a class can be modeled as a Gaussian distribution.[11] The expectation of a voxel belonging to a class can therefore be partially given by the where it falls within the distribution of a classes intensity profile. Solving for parameters such as the intensity variance and mean of a class can therefore result in a more informed,

accurate classification. One example of an EM algorithm used in the tissue classification from MR images of the brain is BRAINSABC[15], which combines probabilities based on intensity with prior spatial probabilities. This method has recently been developed to improve tissue differentiation by including information from a k-nearest neighbor method[16] that encourages voxels within close proximity to each other are more likely to belong to the same class.

Prior knowledge can be intelligently introduced to tissue classification with the use of a Bayesian approach.[17] In these methods, posterior probabilities of a classification are a measure of certainty and are determined by the product of the prior probability and likelihood that is based on new information given by the image. Thus, methods relying on Bayesian analysis can provide superior estimates informed by statistical models for classification as well as intensity models of a given class. The resulting measure of certainty makes the Bayesian approach particularly useful in augmenting EM algorithms.[18]

Level set methodologies have been developed for the segmentation of many different types of images. Level set methods characterize an N-dimensional image with an N+1-dimensional function, of which the zero level defines region boundaries. Typically, a cost function is designed specific to the image model, and includes criteria based on the edge or region information. This can be particularly robust to complex topology in comparison to other methods. Level set segmentation has been proven to complement atlas-based segmentation results[20], and has been applied to the segmentation of neonatal brains in recent works with some success.[21]

1.2.2 Existing Infant Atlases

As a result of the previously mentioned challenges, tissue classification of infant brains from MR images continues to be an open area of research. As the primary chosen method of image segmentation rely on atlas-based algorithms, there are several atlases available intended to represent subjects at various stages in development. Contributing subjects to each atlas were collected locally or taken from institutional image repositories maintained by organizations such as the National Institute for Health.[22,23]

One strategy for generating pediatric atlases with superior template images by utilizing an iterative averaging approach was described by Sanchez et al.[24] Atlases proposed by University of North Carolina[25] and Imperial College London[26] imaging labs are designed for segmentation employing the multi-atlas or subject-specific classification methodology. These methods propose distinct atlases for specific age ranges, typically separated by a year. Another approach retains a collection of scans across a wide age distribution and applies a temporal kernel to weight the contribution of scans in the creation of an individualized atlas for each subject of a given age.[27] As many of these atlases have only been under development for a few year, they typically consist of less than 50 contributing tissue maps.

Unfortunately, each atlas assumes a unique set of scan protocols, tissue classes, and applicable age range. Furthermore, in many cases unmyelinated white is often mislabeled as gray matter due to the gradual contrast reversal observed during development.

1.3 Aims

The goal of this work was to create a framework capable of collecting quantitative volumetric data from images of infant brains for a longitudinal population study investigating developmental differences between subjects born at term and subjects born preterm. Primary subcortical structures as well as gross intracranial tissue types were of interest. Unbiased processing of many subject scans was desired to achieve significant results. To that end, an automated pipeline comprised of open source image analysis tools was developed that included atlas-based segmentation followed by a novel level set segmentation correction. Representative atlases compatible with this pipeline were created from the pool of subject scans available from the study.

Subjects of the study were scanned at discharge of hospital and as part of a follow-up visit at one year of age. This created two subject pools, neonate and one-year-old, for which two age-specific atlases were required. Atlases contain T1 and T2 template images, anatomical landmark fiducials for those template images, and an anatomical landmark model for registration purposes as well as spatial probability maps and intensity priors corresponding to each label of interest for the segmentation algorithm. Ten distinct labels, shown in Table1, were chosen to classify both subcortical structures, cortical structures, and remaining tissue types such as blood and cerebrospinal fluid (CSF).

Table 1: Tissue types included in each atlas
with respective label color value

Neonate	One-year-old	Label Value	Color Map Key
White Matter		1	
Cortical Gray Matter		2	
Unmyelinated White Matter	N/A	3	
CSF		4	
Venous Blood		5	
Cerebellar Gray Matter		11	
Cerebellar White Matter		12	
Basal Ganglia	Corpus Striatum	19	
N/A	Globus Pallidus	23	
Thalamus		24	
Hippocampus		25	

Subject datasets consisted of T1-weighted and T2-weighted structural images collected at 1 mm^{3} isotropic resolution. The desired outputs of the automated pipeline were commonly oriented, bias corrected T1 and T2-weighted images and an accurate label map of the entire intracranial volume.

1.5 Outline of Thesis

The framework of this thesis is as follows:

Section 2. Methods: A thorough description of processes and specifications pertaining to this work is given.

2.0 Current Tools: A brief description of the existing computational image processing tools and methods used throughout this work.

2.1 Automated Pipeline: The individual stages of image processing including preprocessing, initial registration, tissue classification, and a post-processing level set segmentation.

2.2 Generation of Age-Specific Atlases: The process by which each atlas was created. This includes the generation of spatial probability maps, intensity priors, template images, landmark files, and the associated model files required for anatomically driven registration.

2.3 Validation Experiment: A sample of cleaned tissue maps are compared to automated results to determine some measure of the accuracy of the automated pipeline and age-specific atlases.

Section 3. Results: The utility of the proposed tools is examined and the performance of the atlas and automated pipeline are analyzed both quantitatively and qualitatively.

Section 4. Discussion: Concepts and accomplishments are summarized and suggestions for future work are proposed.

2. METHODS

2.1 Current Tools

Much of this work was made possible by the application of many extensively developed image processing tools. With the exception of the level set segmentation, all tools used in the automated pipeline are available within the open software suite BRAINSTools.[28]

2.1.1 BRAINSConstellationDetector

BRAINSConstellationDetector[29] is a fully automated program capable of identifying anatomical landmarks within brain. Models can be trained from a sample of T1 weighted images with respective sets of manually identified landmarks. Landmark templates are created by sampling the images around each landmark. A representative constellation describing the most likely spatial relationship between landmarks is found by solving the linear model from the set of landmark points. The algorithm uses the spatial model to search within a limited region for the area that most closely matches the template for each respective landmark. These automatically identified points can be used to determine an anatomically driven affine registration with a target image for which these landmarks are also known.

The constellation model, landmark weighting for determining affine registration, and set of landmarks for the T1 template image are unique to each age-specific atlas created. The creation of these auxiliary files will be described in a later section.

2.1.2 BRAINSFit

BRAINSFit[30] provides a method of image registration specifically developed for 3D volumes of the brain. Several options of convergence criteria, similarity metrics, and varying strategies of registration methods are available to encourage optimization for a variety of users. One of the methods included is a high-dimensional symmetric image normalization (SyN) registration.[31] SyN registration is ideal for late stage registration of brain images dues to its large capacity for constrained deformation at high resolutions. As implemented by BRAINSFit, SyN registration consists of multiple stages of parametric optimization at increasing resolution.

Within the automated pipeline outlined in this work, BRAINSFit is used to achieve an increasingly accurate step-wise registration between the atlas and subject T1 weighted image during the initial registration prior to tissue classification. The precise parameters of each use will be described in a later section.

2.1.3 BRAINSABC

BRAINSABC[15] is an atlas based method of tissue classification that utilizes an iterative EM algorithm to optimize atlas registration, tissue classification, and bias field correction. The algorithm assigns probabilities for each possible tissue type for a given voxel based on prior spatial and intensity information. A label map is created by selecting the tissue types with the highest likelihoods at each voxel. The parameters chosen for use within this work will be described in a later section.

2.1.4 Level Set Segmentation with Application to MRI

The level set strategy applied in this work was developed by Chunming Li, et. al.[32], and was designed specifically for segmenting images with the types of intensity inhomogeneities found in MR images. It is a region based method that includes a local intensity clustering property that makes it particularly effective in processing noisy images. It is also based on an image model that includes a slowly varying bias field, which is a common complication of image segmentation of MR images.

The cost function that describes the level set in this method is the sum of three criteria and can be iteratively solved to determine the expected intensity of each region, a bias field image, and the region boundary contours. The contour smoothing criterion penalizes arc length and the distance regularization criterion ensures stable convergence depend only on the region contours. The local clustered intensity criterion samples intensity values around each voxel after applying the assumed bias field to determine the relative cost of being within a region and depends on all parameters. Please see Appendix I for a more in depth explanation of each component of the cost function and its solution.

A Matlab[33] implementation of this method for the multiphase segmentation of 2D images was made publically available by the authors. This code was adapted for multiphase segmentation of 3D images to improve differentiation of white and gray matter. This adaptation will be discussed at length in a later section.

2.2 Automated Pipeline

The pipeline developed in this work can be reduced to four stages: preprocessing, initial atlas-to-subject registration, atlas based tissue classification, and post-processing

correction of labels using level set segmentation. This pipeline requires an input of T1 and T2 weighted images and the age-appropriate atlas. It produces a label map of the entire intracranial volume and AC-PC aligned bias corrected T1 and T2 weighted images. Figure 2 presents a flow chart that displays the direction of inputs and outputs at each stage of the pipeline.

2.2.1 Preprocessing

To ensure a degree of regularity across datasets, a number of automated preprocessing steps were completed before performing registration and segmentation. The T1 weighted image of each subject was aligned in AC-PC space, cropped at a uniform distance below the anterior commissure to reduce unnecessary computations, and then processed to automatically identify a set of anatomical landmarks. This was done using BRAINSConstellationDetector. The respective T2 weighted image was rigidly registered to the T1 weighted image using BRAINSFit and Mattes Mutual Information as a metric. Finally, N4 bias correction[34] was applied to both the T1 and T2 weighted images.

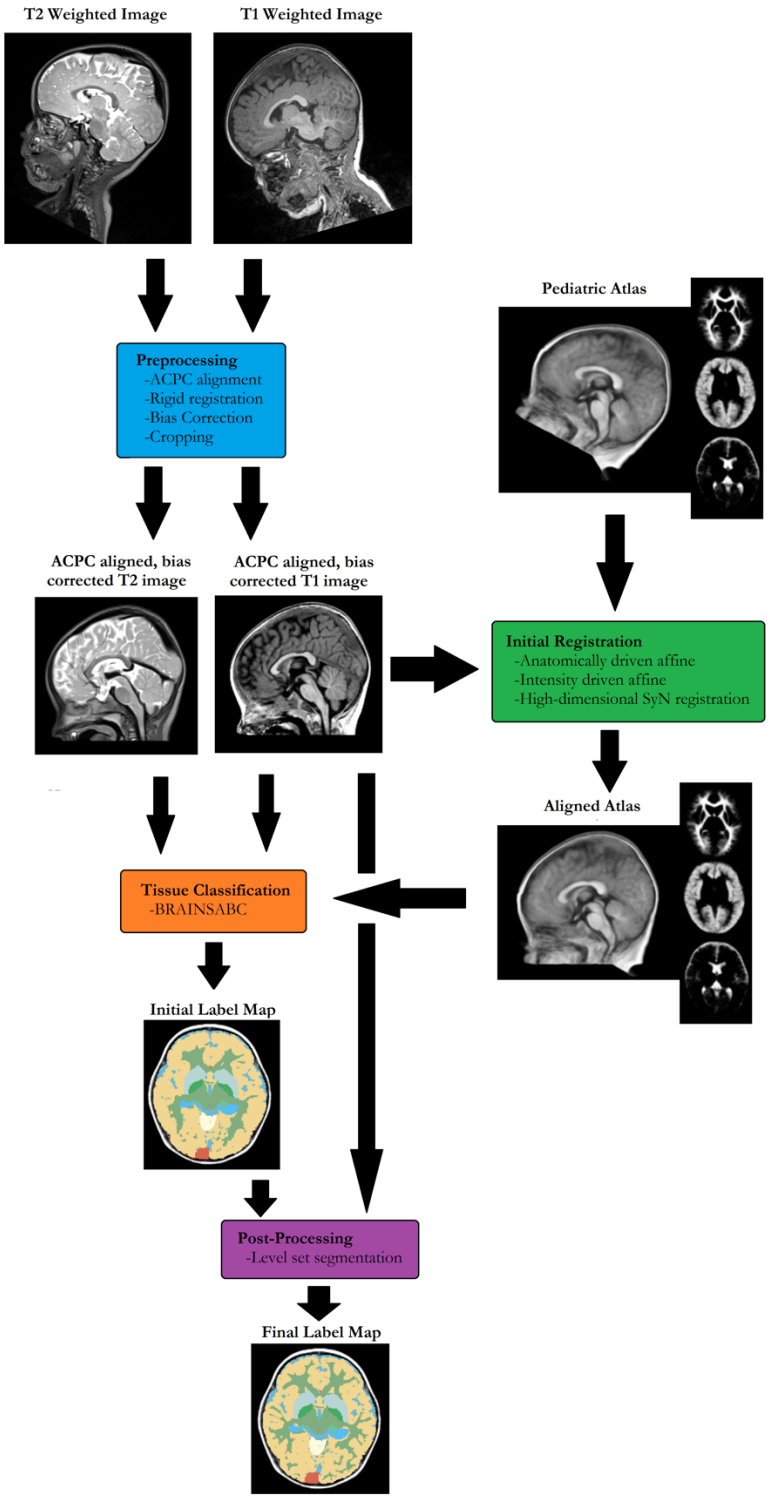

Figure 2: A flow chart describing the inputs and outputs of each stage in the automated pipeline.

2.2.2 Initial Registration

After preprocessing, the atlas was incrementally registered to the subject using Mattes Mutual information as a metric and the T1 template image provided by the atlas. The first step performed was solving for the affine transformation found by comparing the atlas landmarks to the subject landmarks. This least squares problem is solved by BRAINSConstellationDetector and was weighted with respect to the importance and expected accuracy of each landmark. This transform is used to initial an intensity driven affine registration which is performed by BRAINSFit. Finally, the resulting transform was used to initialized a multiresolution SyN registration, again using BRAINSFit. Due to the high dimensionality of SyN registration, this step is the most computationally expensive portion of the pipeline. To avoid errors due to resampling, and for convenience, the concatenated transform is passed to the BRAINSABC. Figure 2 demonstrates the increasing accuracy of each step in the atlas-to-subject initial registration.

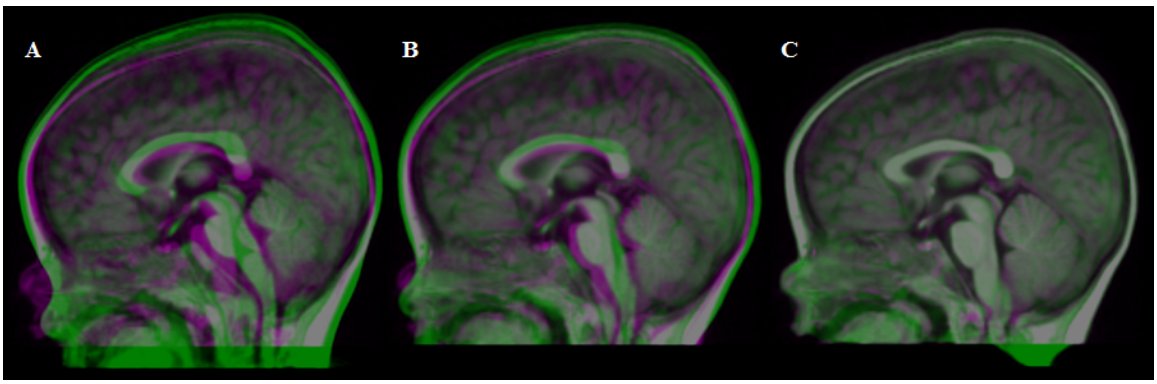

Figure 3: Each step in initial registration: AC-PC alignment (A), anatomically driven affine transformation (B), and high-dimensional SyN Registration (C). The atlas is visualized in green and the subject in visualized in magenta.

2.2.3 Tissue Classification

Initial tissue classification was generated by BRAINSABC. This required the initial atlas-to-subject transform given by the initial registration, the coregistered T1 and T2 weighted subject images, and the age specific atlas consisting of spatial and intensity priors. As an initial bias correction had already been performed in preprocessing, bias correction capabilities of the iterative method in BRAINSABC were considerably restricted. Results were required less than ten iterations until no further improvement was detectable. The output of BRAINSABC included a final atlas-to-subject registration, further bias corrected T1 and T2 weighted images, and a label map classifying the intracranial volume into ten tissue types.

While subcortical structures were segmented with sufficient accuracy using age-appropriate atlases with a relatively small contributing subject pool, cortical white and gray matter differentiation suffered greatly. Therefore, a secondary segmentation targeted specifically for this tissue border was developed.

2.2.4 Level Set Method

A MATLAB[33] implementation of the level set strategy described previously for the segmentation of 2D MR images was made publically available. This included a multiphase approach which segmented the image into three distinct regions. In order to feasibly apply this method for the segmentation of a 3D image, a few important changes needed be made. The Neumann Boundary condition needed to be satisfied to avoid divergence at the edges of the image. The contour smoothing criterion needed to be converted from an arc length penalty to a surface curvature penalty.

Most notably, every iteration in the 2D method required the calculation of four convolutions of the K-means windowed kernels with various intermediate images of similar size to the observed image. This was computationally practical for images of size 2^8 by 2^8 pixels and kernels of size roughly 2^4 by 2^4 pixels. When a third dimension of similar length with respect to each construct is considered, the computational cost of each iteration easily exceeds practicality. To address this, convolution was substituted for element by element multiplication in the frequency domain. For further efficiency, images were cropped to the bounding box given by the label map produced by BRAINSABC and intermediate images were passed in the frequency domain when convenient. To prevent aliasing artifacts, images were heavily zero padded.

Next, the method was optimized for cortical tissue differentiation for each age group. For the one-year-old subjects, this meant establishing parameters that could reliably correct mislabeled cortical gray matter and white matter. For the neonate subjects the target labels were cortical gray matter and unmyelinated white matter. Because the T1 weighted image presented with the best contrast between white and gray matter for one-year-old subjects, it was chosen as the input. The neonates suffered from very poor SNR and edge definition. To address this, the difference image of the normalized T1 and T2 weighted images was used to enhance cortical gray matter and unmyelinated white matter contrast. An example of this image is shown in Figure 4. For the one-year-old group, a K-means kernel designed as a 3D Gaussian window with sigma of 3 mm and windowing radius of 7 mm gave the best results. Due to the smoothness already present in the neonate input image, a similarly designed kernel with sigma of 2 mm and windowing radius of 3 mm yielded the best differentiation between cortical tissues.

The input images from both age groups were saturated at both intensity extremes to encourage distinction of regions within an intensity range specific to cortical tissues. Several of the tunable parameters performed optimally for both age groups. The weighting of the cost function criteria, the time step used in during the gradient descent solution of the level set contours, and the choice to distance regularization function were chosen to be consistent with the recommendation by the literature.[32]

To foster the most direct convergence, the method was initialized with regions given by the labels resulting from BRAINSABC and the intensity values of each region were set to expected values for white matter, gray matter, etc. for the first iteration. Sufficient convergence and satisfactory results were found to be reliably met in less than 30 iterations.

The regions generated from this method were used to correct the target labels to provide more accurate cortical tissue classification. However, the spatial priors used in BRAINSABC allowed for a better performance of tissue classification compared to this method when partial volume effect was a significant factor. To play to the strengths of both methods, only voxels within each target label that fell outside of a mask created by dilating the existing CSF labels were considered for correction. Results showing cortical tissue differentiation before and after this method are displayed in Figure 4.

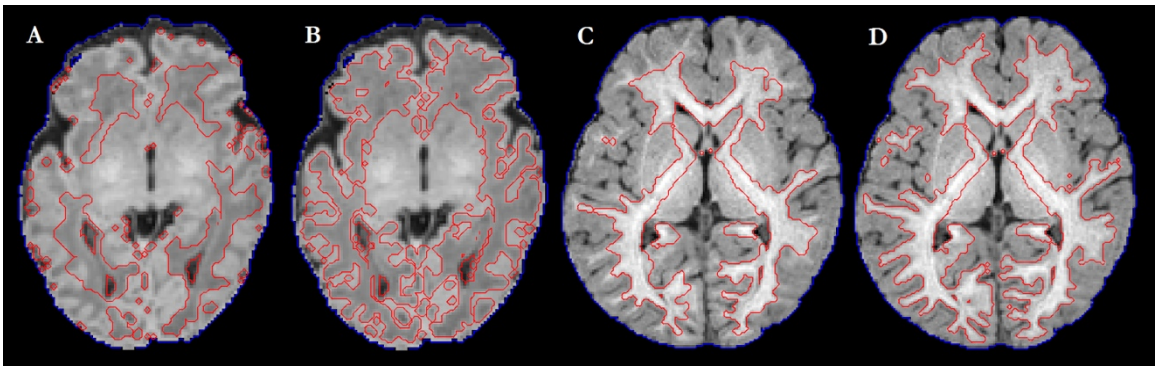

Figure 4: White/gray matter separation before (A,C) and after (B,D) level set segmentation on a neonatal subject (A,B) and a one-year-old subject (C,D). The image shown here for the neonate scan is the masked difference image (normalized T1-T2) while a T1-weighted image is shown for the one-year-old.

2.3 Generation of Age-Specific Atlases

The previously mentioned pipeline requires several components based on prior knowledge of the subject images. These include spatial probability maps derived from age appropriate tissue maps, a T1 and T2 weighted template image for the registration required to apply those spatial probability maps, and intensity distributions for each tissue type based on sampling tissue intensities from similarly aged subjects. Ten tissue types were chosen as labels in the one-year-old atlas including: cerebrospinal fluid, cortical gray matter, white matter, corpus striatum, globus pallidus, thalamus, hippocampus, cerebellar white matter, cerebellar gray matter, and venous blood. The neonate atlas included an additional label for unmyelinated white matter, and merged the globus pallidus label and the corpus striatum labels to form a "basal ganglia" label due to a lack of a consistently clear differentiation between the globus pallidus and the putamen amongst the neonate scans. Also included in an atlas compatible with this pipeline are the landmark model and weighting required for anatomically driven registration and the landmarks associated with the template images.

The subjects chosen to contribute to each atlas were taken from the respective age groups that were under investigation and were subjected to the preprocessing done in the automated pipeline. The T1 template image was created by averaging all T1 images with contributing cleaned tissue maps in a common coordinate space after high dimensional registration to an arbitrary contributing subject. The T2 images averaged to create the T2 template image after each was also resampled to this coordinate space. An upper and lower quartile for the intensity distributions of each tissue type was determined by sampling both T1 and T2 weighted images from several subjects at various locations within each tissue type. These intensity values were purposefully narrowed during atlas creation to force a more intensity driven segmentation while spatial priors were poor, and were broadened as spatial priors improved to provide a segmentation driven by a more balanced contribution from spatial and intensity priors.

The spatial probability maps require a collection of accurate tissue maps from a group of images that is representative of the subjects which will be segmented. The tissue maps are transformed to the common coordinate space by registering each respective T1 weighted image with the T1 template image in a similar manner used in the initial registration of the automated pipeline. Then each label is individually averaged and Gaussian smoothed to provide a value from zero to one at each voxel representing the likelihood of that voxel belonging each label based on its spatial location. Figure 3 displays an axial slice of the spatial probability maps for several tissue types.

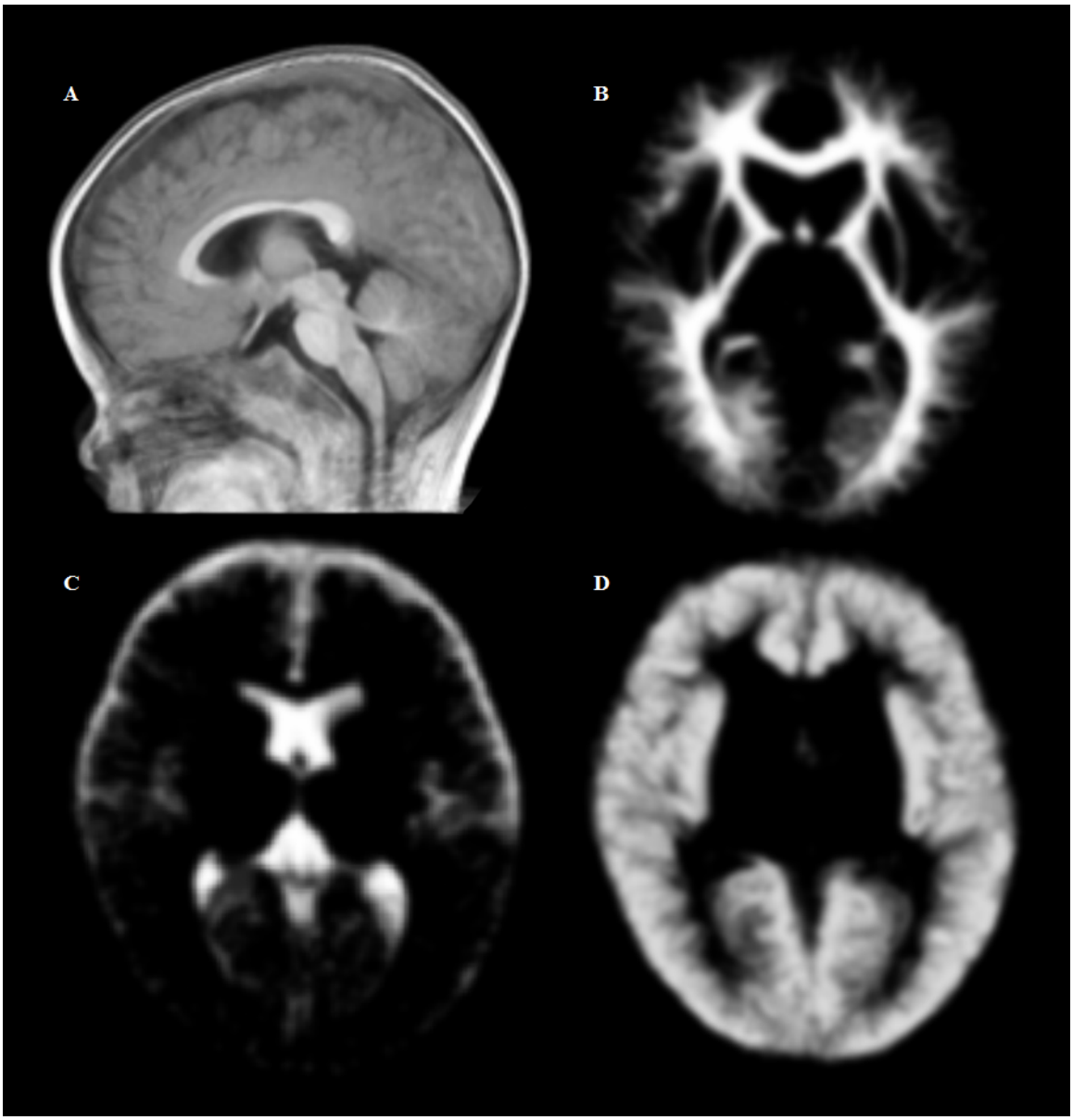

Figure 5: Individual components of the one-year-old atlas: a sagittal slice of the T1 template image (A), and axial views of the spatial priors for white matter (B), CSF (C), and cortical gray matter (D).

In order to limit the necessity of manual intervention in creating these tissue maps, the results of the automated pipeline and an insufficient atlas were manually cleaned and incrementally added to the atlas until representative spatial probability maps were realized. To obtain the first few tissue maps of the one-year-old atlas an existing adult atlas was

applied to one-year-old subject datasets with the pipeline. After manual cleaning, an atlas derived from just two subjects greatly outperformed the adult atlas. This process was repeated until the one-year-old atlas was sufficiently representative. The neonate atlas was created similarly, using the one-year-old priors as the initial atlas.

To determine the linear model utilized by BRAINSConstellationDetector to search for anatomical landmarks, a collection of manually selected landmarks of T1 weighted images is required. Ten subjects from each age group were randomly selected to create the model. Twenty landmarks spread throughout the intracranial volume were chosen and identified on each subject. Given the spatial relationship of the landmarks, a linear model was solved using ancillary BRAINSTools programs. The anatomical landmarks of the T1 and T2 template images were given by the application of this model and BRAINSConstellationDetector to the contributing subject that was chosen in determining the common coordinate space. The weighting of the landmarks for affine registration was based on the the accuracy and reliability of each landmarks identification.

2.4 Validation Experiment

To determine a quantitative measure for the accuracy of the automated pipeline and the level set segmentation, the results were compared to manually corrected tissue maps using Dice coefficients, Hausdorff distances, and Intraclass Correlation coefficients. Expert manual corrections were performed on three slices at orthogonal orientations: axial at the anterior commissure, coronal at the posterior commissure, and sagittally intersecting the hippocampus. Samples from two subjects in each age pool were investigated.

Subjects in the one-year-old pool ranged in age between 10 to 18 months at the time of the scan. The neonate pool was comprised of subjects born preterm and imaged at the time of discharge from the hospital. The age for the neonate subjects ranged from zero to three months corrected age. Both the neonate and one-year-old atlases were created from eight datasets with manually inspected tissue maps. Scanning was done with a 3T Siemens TIM Trio scanner. Images from a T1-weighted magnetization-prepared rapid-acquisition gradient-echo (MP-RAGE) sequence and a turbo spin-echo T2-weighted sequence were acquired. Both scans were taken in the coronal plane at 1mm isotropic resolution.

3. RESULTS

Accurate tissue classification of several subcortical structures and gross intracranial tissue types was achieved. Segmentation results benefited from multimodal input and anatomically driven registration. Tissue classification visually improved with the addition of each component in the segmentation pipeline. The variance of the Gaussian filtering after averaging tissue maps in a common coordinate space during atlas creation was minimized for optimal subcortical structure segmentation, and the level set segmentation drastically improved cortical white and gray matter differentiation.

As reported in Table 2, the level set segmentation greatly increased the accuracy of cortical white and gray matter labels in both the neonatal and one-year-old subjects. This, in turn, considerably improved the overall accuracy of the segmentation. The higher Dice and ICC coefficients suggest improved similarity and correlation between cleaned label maps and the final results of the pipeline. As reported in Table 3, the reduced average Hausdorff distance from the manually cleaned labels after the level set correction

Table 2: The Dice and ICC coefficients when comparing atlas and level set segmentation results to manually cleaned classifications. *The slices used in the validation experiment did not contain enough cerebral white matter voxel to calculate significant values.

Validation Results		**Before Level Set**		**After Level Set**	
		Dice	**ICC**	**Dice**	**ICC**
Neonate	Unmyelinated White Matter	**0.608**	**0.599**	**0.953**	**0.952**
	Gray Matter	**0.697**	**0.688**	**0.919**	**0.920**
	White Matter	**0.880**	**0.855**	**---**	**---**
	CSF	**0.908**	**0.907**	**---**	**---**
	Venous Blood	**0.900**	**0.888**	**---**	**---**
	Cerebellar GM	**0.751**	**0.902**	**---**	**---**
	Cerebellar WM	**N/A***	**N/A***	**---**	**---**
	Basal Ganglia	**0.896**	**0.899**	**---**	**---**
	Thalamus	**0.800**	**0.829**	**---**	**---**
	Hippocampus	**0.860**	**0.841**	**---**	**---**
	All Labels	**0.732**	**0.880**	**0.916**	**0.882**
One-Year-Old	White Matter	**0.847**	**0.840**	**0.983**	**0.982**
	Gray Matter	**0.911**	**0.903**	**0.984**	**0.983**
	CSF	**0.991**	**0.991**	**---**	**---**
	Venous Blood	**0.887**	**0.902**	**---**	**---**
	Cerebellar GM	**0.965**	**0.987**	**---**	**---**
	Cerebellar WM	**N/A***	**N/A***	**---**	**---**
	Striatum	**0.974**	**0.965**	**---**	**---**
	Thalamus	**0.959**	**0.982**	**---**	**---**
	Hippocampus	**0.924**	**0.920**	**---**	**---**
	Globus Pallidus	**0.990**	**0.977**	**---**	**---**
	All Labels	**0.911**	**0.971**	**0.982**	**0.972**

verify a more accurate delineation of white and gray matter. In general, the one-year-old subjects were more accurately segmented than the neonatal subjects. This can be attributed to a lower attainable SNR with the younger subjects.

Table 3: The average Hausdorff distance to the manually cleaned labels.

Average Hausdorff Distance		**Before Level Set Correction**	**After Level Set Correction**	**Change**
Neonate	**Unmyelinated White Matter**	**1.01 mm**	**0.12 mm**	**-0.89 mm**
	Gray Matter	**0.49 mm**	**0.21 mm**	**-0.28 mm**
One-Year-Old	**White Matter**	**0.49 mm**	**0.05 mm**	**-0.44 mm**
	Gray Matter	**0.12 mm**	**0.03 mm**	**-0.09 mm**

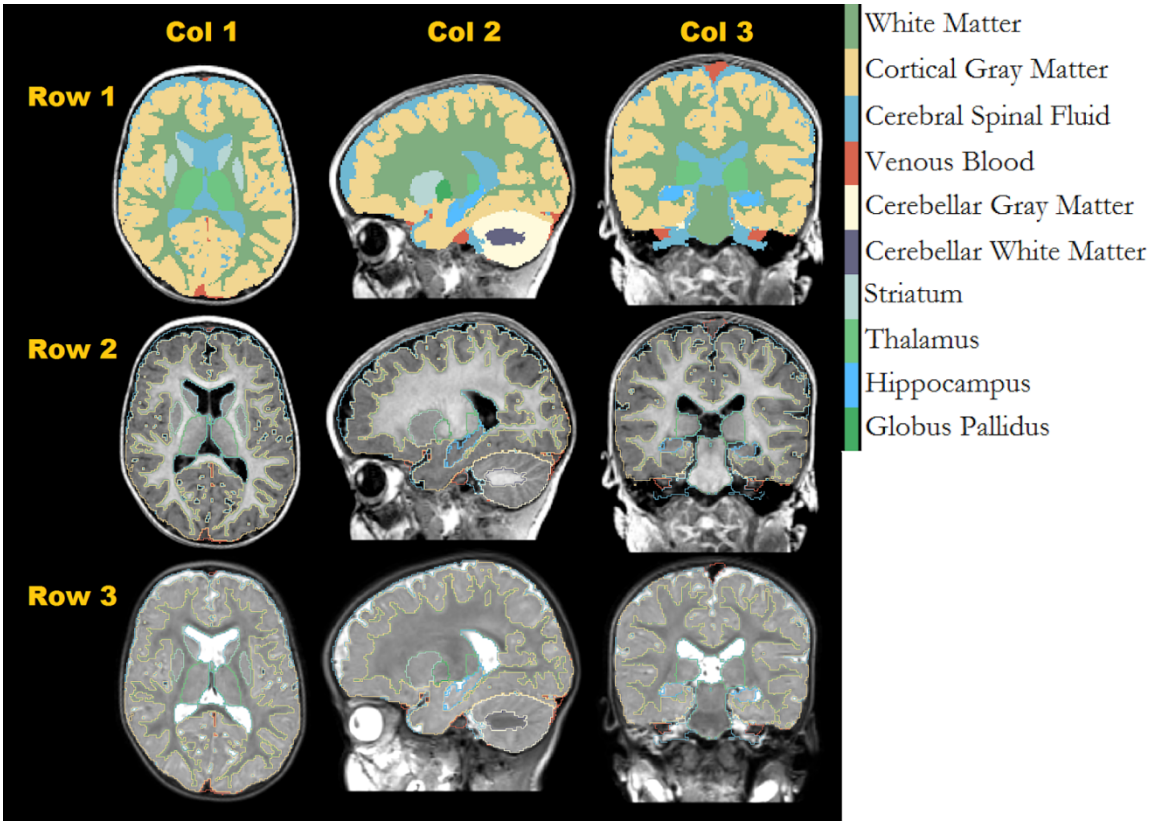

Figure 6: Final classification results of a typical subject using the one-year-old atlas. Row 1 shows the label values, Row 2 plots label borders on the T1 bias corrected image, and Row 3 plots the label borders on the T2 bias corrected image.

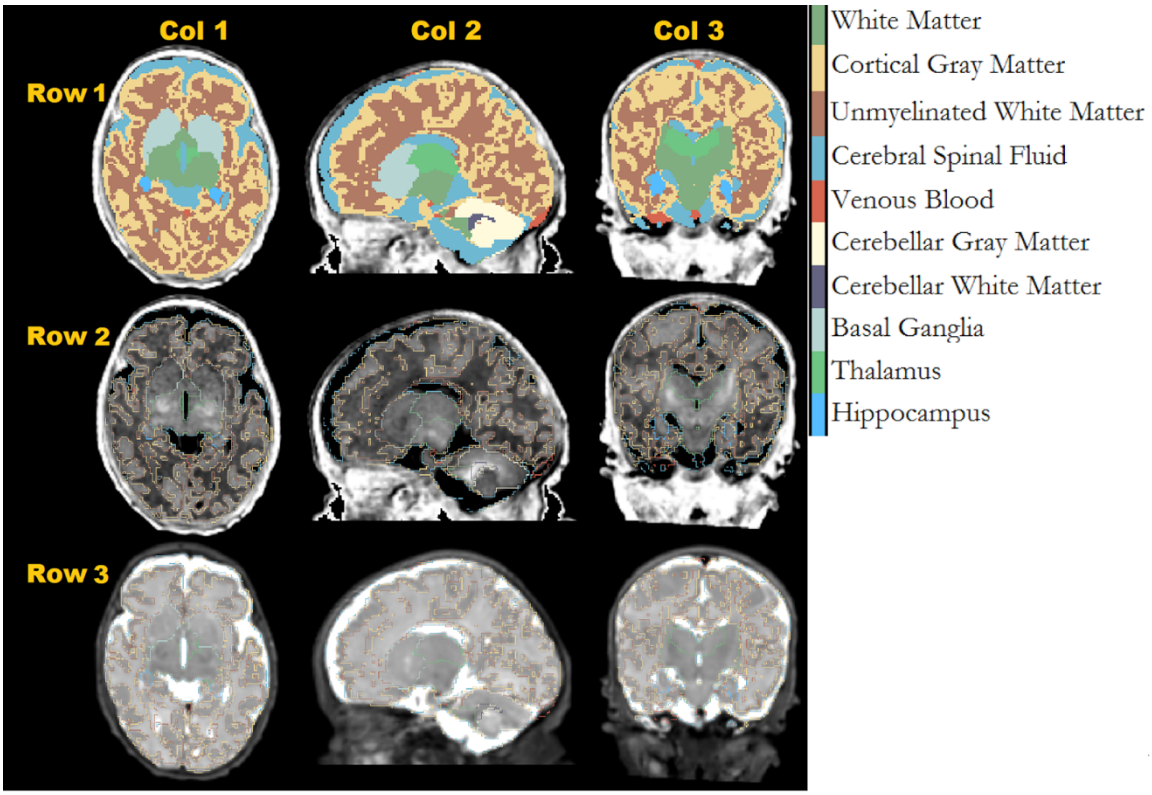

Figure 7: Final segmentation results on a typical subject using the neonate atlas. Column 1 gives equally positioned axial slices; Column 2, sagittal slices; and Column 3, coronal slices.

4. DISCUSSION

An early version of the one-year-old atlas was used for volumetric analysis and provided significant results in a study comparing neurodevelopment of infants born pre- and full-term.[35] In said work, Talairach boxes[3] were used to isolate functionally distinct regions for tissue volume comparisons in conjunction with cognitive analysis. Future work may attempt to improve the accuracy, expand on the capabilities of this atlas, or apply these tools to a similar pediatric neurological condition.

Due to the particularly high spatial variation present in pediatric subjects in this age range, the accuracy of the initial atlas-to-subject registration weighs heavily on the accuracy of the final results. This is why the multiresolution high-dimensional SyN

registration is afforded the greatest proportion of processing time compared to any of the other steps in the pipeline. Any improvements that can be made to the efficiency or accuracy of this step should be held as a high priority.

The addition of multimodal information could also improve the accuracy of the initial registration as well as tissue classification. One such area that would benefit in particular from this addition would be the separation between the cerebellum and the occipital lobe. This border is markedly noticeable in T2 weighted images due to the modality-specific contrast between the venous blood of the transverse sinuses compared to cortical gray matter and cerebellar gray matter. If multimodal information was utilized in the registration of the atlas to the subject, the spatial priors would be applied more accurately in this area. To benefit classification, the level set method could be further developed to process multimodal input informed by Bloch equation simulations in order to provide improved differentiation of several additional tissue types.

Classification can also be improved with a more representative set of prior knowledge. As with all atlases, priors can be made increasingly representative by increasing the sample size of contributing subjects. This is also true for the landmark model, which would benefit from contributions from a greater number of subjects with identified landmarks as well as a selective expansion of the anatomical landmarks included in the model. This would further address the accuracy and efficiency of initial registration.

This work has created a framework for several further analyses that might provide useful insight. One such approach is the increased parcellation of labels by separating functionally specific regions of each current label. For example, it is not uncommon for

an adult atlas to include hundreds of labels corresponding to specific gyri or lobes. This will be in part limited by the SNR and image resolution obtainable by the pulse sequences. Fiber tracking, diffusion, and surface analysis are also potential avenues of continuation for this work.

While this atlas was created for and comprised of scans of pre- and full term infants, the automated pipeline and approach to atlas generation aided by similar spatial probabilities in this work could be applied to a number of pediatric neurological conditions. A few potential disease targets with particularly high impact include autism and cerebral palsy.[36]

APPENDIX

All equations and reasoning taken from IEEE Trans Image Process **20,** pg 2007–2016.[32]

This method assumes a noisy piecewise constant model with N distinct regions affected by a slowly varying bias field. This gives us Eqn. A.1, where *J* is the true image, *b* is the bias field, *n* is noise, and *I* is the observed image.

Eqn. A.1) $$I = bJ + n$$

For an image with two regions, the intensity value of each region is given by the set $c = \{c_1, c_2\}$, and the region boundary is ϕ. To identify each region by solving for ϕ, the cost (*F*) function given by Eqn. A.2 is proposed.

Eqn. A.2) $$F(\phi, c, b) = E(\phi, c, b) + \upsilon L(\phi) + \mu R(\phi)$$

E is a K-means clustering criterion given by Eqn. A.3, *L* is a contour smoothing criterion given by Eqn. A.4, and *R* is a distance regularization term given by Eqn. A.5. The variables μ and υ are criterion weighting factors.

Eqn. A.3) $$E(\phi, c, b) = \int \sum_{i=1}^{N} (\int K(y - x)|I(x) - b(y)c_i|^2 dy) M_i(\phi(x)) dx$$

K is a windowed smoothing kernel. A Gaussian filter with a fixed σ was used in this work. M_i is the region associated with the intensity value c_i.

Eqn. A.4) $$L(\phi) = \int |\nabla H(\phi)| dx$$

H is the Heaviside function which converts the contour to distinct regions.

Eqn. A.5) $$R(\phi) = \int P(|\nabla \phi|) dx$$

P is the signed distance function.

This requires iteratively solving for *c*, *b*, and ϕ. The solutions for each are given by Eqn. A.6, A.7, and A.8 respectively. Note, that gradient descent must be used to solve for ϕ.

Eqn. A.6) $$c_i = \frac{\int (b * K) I M_i(\phi(y)) dy}{\int (b^2 * K) M_i(\phi(y)) dy}, \quad i = 1, \ldots, N$$

Eqn. A.7) $$b = \frac{(I[\sum_{i=1}^{N} c_i M_i(\phi(y))]) * K}{[\sum_{i=1}^{N} c_i^2 M_i(\phi(y))] * K}$$

Eqn. A.8) $$\frac{\partial \phi}{\partial t} = \frac{\partial F(\phi, c, b)}{\partial \phi}$$ (see pg. 2011[32] for a more explicit eqn.)

BIBLIOGRAPHY

1. Dobbing, J. & Sands, J. Quantitative growth and development of human brain. *Arch Child* **48,** 757–767 (1973).

2. Knickmeyer, R. C. *et al.* A structural MRI study of human brain development from birth to 2 years. *J Neurosci* **28,** 12176–12182 (2008).

3. Talairach, J. & Tournoux, P. *Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging*. (Thieme, 1988).

4. Williams, L.-A. *et al.* Optimization of 3D MP-RAGE for neonatal brain imaging at 3.0 T. *Magn Reson Imaging* **25,** 1162–1170 (2007).

5. Conklin, J., Winter, J. D., Thompson, R. T. & Gelman, N. High-contrast 3D neonatal brain imaging with combined T1- and T2-weighted MP-RAGE. *Magn Reson Med* **59,** 1190–1196 (2008).

6. Gilmore, J. H. *et al.* Regional gray matter growth, sexual dimorphism, and cerebral asymmetry in the neonatal brain. *J. Neurosci.* **27,** 1255–1260 (2007).

7. Huang, H. *et al.* White and gray matter development in human fetal, newborn and pediatric brains. *Neuroimage* **33,** 27–38 (2006).

8. Otsu, N. A threshold selection method from gray-level histograms. *Automatica* **11,** 23–27 (1975).

9. Abo-Eleneen, Z. Thresholding based on Fisher linear discriminant. *J. Pattern Recognit. Res.* **2,** 326–334 (2011).

10. Suzuki, H. & Toriwaki, J. Automatic segmentation of head MRI images by knowledge guided thresholding. *Comput. Med. Imaging Graph.* **15,** 233–240 (1991).

11. Pham, D. L., Xu, C. & Prince, J. L. Current methods in medical image segmentation 1. *Annu. Rev. Biomed. Eng.* **2,** 315–337 (2000).

12. Egmont-Petersen, M., de Ridder, D. & Handels, H. Image processing with neural networks—a review. *Pattern Recognit.* **35,** 2279–2301 (2002).

13. Lu, S. W. & Xu, H. Textured image segmentation using autoregressive model and artificial neural network. *Pattern Recognit.* **28,** 1807–1817 (1995).

14. Gelenbe, E., Feng, Y. & Krishnan, K. R. R. Neural network methods for volumetric magnetic resonance imaging of the human brain. *Proc. IEEE* **84,** 1488–1496 (1996).

15. Young Kim, E. & Johnson, H. J. Robust multi-site MR data processing: iterative optimization of bias correction, tissue classification, and registration. *Front Neuroinform* **7,** 29 (2013).

16. Ghayoor, A., Paulsen, J. S., Kim, R. E. & Johnson, H. J. Tissue classification of large-scale multi-site MR data using fuzzy k-nearest neighbor method. in *SPIE Medical Imaging* 97841V–97841V (International Society for Optics and Photonics, 2016).

17. Hanson, K. M. Introduction to Bayesian image analysis. in *Medical Imaging 1993* 716–731 (International Society for Optics and Photonics, 1993).

18. Marroquín, J. L., Vemuri, B. C., Botello, S., Calderón, F. & Fernandez-Bouzas, A. An accurate and efficient Bayesian method for automatic segmentation of brain MRI. *Med. Imaging IEEE Trans. On* **21,** 934–945 (2002).

19. Cabezas, M., Oliver, A., Lladó, X., Freixenet, J. & Cuadra, M. B. A review of atlas-based segmentation for magnetic resonance brain images. *Comput Methods Programs Biomed* **104,** e158–77 (2011).

20. Duay, V., Houhou, N. & Thiran, J. P. Atlas-based segmentation of medical images locally constrained by level sets. in *Image Processing, 2005. ICIP 2005. IEEE International Conference on* **2,** II–1286–9 (ieeexplore.ieee.org, 2005).

21. Wang, L. *et al.* Segmentation of neonatal brain MR images using patch-driven level sets. *Neuroimage* **84,** 141–158 (2014).

22. Job, D. E. *et al.* A brain imaging repository of normal structural MRI across the life course: Brain Images of Normal Subjects (BRAINS). *NeuroImage* (2016).

23. Evans, A. C. The {NIH} {MRI} study of normal brain development. *NeuroImage* **30,** 184 – 202 (2006).

24. Sanchez, C. E., Richards, J. E. & Almli, C. R. Neurodevelopmental MRI brain templates for children from 2 weeks to 4 years of age. *Dev. Psychobiol.* **54,** 77–91 (2012).

25. Cherel, M. *et al.* Automatic tissue segmentation of neonate brain MR Images with subject-specific atlases. in *SPIE Medical Imaging* 941311–941311–11 (International Society for Optics and Photonics, 2015).

26. Gousias, I. S., Hammers, A., Counsell, S. J., Srinivasan, L. & others. Magnetic resonance imaging of the newborn brain: automatic segmentation of brain images into 50 anatomical regions. *PLoS One* (2013).

27. Serag, A. *et al.* Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression. *Neuroimage* **59,** 2255–2265 (2012).

28. Kim, R. & Johnson, H. Efficient and Extensible Workflow: Reliable whole brain segmentation for Large-scale, Multi-center Longitudinal Human MRI Analysis using High Performance/Throughput Computing Resources. (2015).

29. Ghayoor, A., Vaidya, J. G. & Johnson, H. J. Development of a novel constellation based landmark detection algorithm. in *SPIE Medical Imaging* 86693F–86693F–6 (International Society for Optics and Photonics, 2013).

30. Johnson, H., Harris, G., Williams, K. & Others. BRAINSFit: mutual information rigid registrations of whole-brain 3D images, using the insight toolkit. *Insight J* 1–10 (2007).

31. Avants, B. B. *et al.* A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* **54,** 2033–2044 (2011).

32. Li, C. *et al.* A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE Trans Image Process* **20,** 2007–2016 (2011).

33. *MATLAB*. (The MathWorks Inc., 2015).

34. Tustison, N. J. *et al.* N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* **29,** 1310–1320 (2010).

35. Benavides, A. Early Neurodevelopmental Outcomes in Preterm Infants: Memory, Attention, and Encoding Speed. (University of Iowa, 2015).

36. Hirtz, D. *et al.* How common are the 'common' neurologic disorders? *Neurology* **68,** 326–337 (2007).