Atlas of Classifiers for Brain MRI Segmentation

B. Kodner[1,2], S. H. Gordon[1,2], J. Goldberger[3] and T. Riklin Raviv[1,2]

[1]Department of Electrical and Computer Engineering, [2]The Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Israel
[3]Faculty of Engineering, Bar-Ilan University, Israel

Abstract. We present a conceptually novel framework for brain tissue segmentation based on an *Atlas of Classifiers* (AoC). The AoC allows a statistical summary of the annotated datasets taking into account both the imaging data and the corresponding labels. It is therefore more informative than the classical probabilistic atlas and more economical than the popular multi-atlas approaches, which require large memory consumption and high computational complexity for each segmentation. Specifically, we consider an AoC as a spatial map of voxel-wise multinomial logistic regression (LR) functions learned from the labeled data. Upon convergence, the resulting fixed LR weights (a few for each voxel) represent the training dataset, which might be huge. Segmentation of a new image is therefore immediate and only requires the calculation of the LR outputs based on the respective voxel-wise features. Moreover, the AoC construction is independent of the test images, providing the flexibility to train it on the available labeled data and use it for the segmentation of images from different datasets and modalities.
The proposed method has been applied to publicly available datasets for the segmentation of brain MRI tissues and is shown to outreach commonly used methods. Promising results were obtained also for multimodal, cross-modality MRI segmentation.

1 Introduction

The prevalence of Magnetic Resonance Imaging (MRI) nowadays for clinical purposes and research contributes to an exponential growth in MRI datasets available. This in turn, accelerates the development of automatic and semi-automatic tools for image analysis and brain tissue segmentation in particular. Despite over two decades of efforts, brain parcellation into gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) remains a difficult problem due to low SNR, partial volume effect, bias field and other imaging artifacts. Additional information, mainly spatial, is therefore required, where image intensities and gradients are not sufficiently discriminative.

Bayesian approaches for brain MRI segmentation [1, 3, 7, 16] are based on the observation that the image intensity distribution can be modeled, with good approximation, by a mixture of Gaussians. Prior probabilities are usually obtained by averaging aligned manual annotations. These priors, also called probabilistic

atlases, define the spatial probability that a voxel at that particular location belongs to a particular brain tissue.

While the classical atlas-based methods have shown to yield successful segmentation results and are applied in commonly used tools, such as FreeSurfer [3], SPM [1], Fast-FSL [18] and Slicer EM segmenter [10], a main concern is that a single atlas may not capture the large heterogeneity of structural brain images. To address this issue the multi-atlas segmentation (MAS) approach emerged, see [6] and references therein. The fused atlas is specific to a target image as it is based on its similarity to each image (or image patches) in the database. Numerous MAS methods were proposed, based on different global, local and non-local similarity measures and a variety of fusion techniques. The flexibility of adapting the atlas to the test images comes at the price of large memory storage, keeping an entire database along with the associated labels, and high computational complexity as the pairwise similarities between a test image and each of the training instances, or (more economically) clusters in the training set [14], are considered.

Machine learning (ML) approaches for segmentation, e.g. Support Vector Machines (SVM), Random Forests (RF) [15, 17] and recently deep neural networks (DNN), e.g. [9, 19] use comprehensive set of features to discriminate between the different tissues. These methods provide promising segmentation results, yet at the cost of long and computationally expensive training process, which requires either huge annotated datasets or smart data augmentation.

We hereby propose a conceptually new form of atlas termed an *atlas of classifiers* (AoC), which enables the abstraction of possibly huge datasets with very few parameters for each voxel. Each classifier is a multinomial logistic regression (LR) function. The weights of the LR functions are determined during training (based on the available labeled data) via gradient ascent processes that aim to maximize a regularized objective function. Once the training phase is completed the classifiers' parameters remain fixed. Multi-class soft segmentation of a new image is performed promptly, by simply applying to each of the image voxels the associated multinomial LR function. Note that unlike common ML approaches, the classifiers are voxel-specific and not class-specific. Solving large number of voxel-wise optimization problems with very few parameters (features) each, significantly facilitates and accelerates the entire training process.

In this paper, the partition of brain MRIs into tissues is considered, yet, the proposed paradigm is general and not tailored to a specific application. To allow multi-modal and cross modality segmentation, rather than using the image intensities directly, each voxel is associated with an intensity-based feature vector. These voxel-wise vectors are key components both in the segmentation process, as inputs to the objective function of the LR weights, and for the alignment of the training data into a common-space for the generation of the atlas.

The rest of the paper is organized as follows. In Section 2 we present the atlas of classifiers. Promising experimental results for IBSR18 and IBSR20 [11] as well as cross-modality segmentation for MRBrainS13 multi-modal data are presented in Section 3. We conclude in Section 4.

2 The Atlas of Classifier Approach

Let $\{I_n\}_{n=1}^N$, where $I_n \colon \Omega \to \mathbb{R}^D$ denote N unimodal ($D = 3$) or multimodal ($D = 4$) training images. Let $\mathcal{I}_n \triangleq \{f^m(I_n)\}_{m=1}^M$, define M features or filters associated with I_n. We denote by $\mathcal{Y}_1, \ldots, \mathcal{Y}_N$ the corresponding labeling functions such that $\mathcal{Y}_n \colon \Omega \to \{1, \ldots, L\}$ partitions the image domain of I_n into L disjoint regions of interest (ROIs), $\Omega_1, \ldots, \Omega_L$. Our goal can be formulated as follows: given a labeled training set, find the segmentation of a test image I_{TEST} by solving the following maximum a posteriori (MAP) problem:

$$\hat{\mathcal{Y}}_{\text{TEST}} = \arg\max_{\mathcal{Y}_{\text{TEST}}} p(\mathcal{Y}_{\text{TEST}} | \mathcal{I}_{\text{TEST}}, \mathcal{I}_1, \ldots, \mathcal{I}_N, \mathcal{Y}_1, \ldots, \mathcal{Y}_N). \quad (1)$$

Let V represents the 3D image size. The proposed AoC allows a statistical summary of the training sets represented by $V \times M \times L$ parameters, regardless of the number of scanned subjects ($N >> M$) and modalities. For comparison, the multi-atlas approach requires voxel-wise storage of $V \times N \times L$ for unimodal data.

2.1 AoC Model Overview

The AoC concept is illustrated in Fig. 1. The training phase (light blue background) is performed once for a given database of MRI scans and the respective label maps. Each of the raw images is mapped into a feature space which, together with the associated labels, contributes to the generation of the voxel-wise LR functions that represent the atlas, see Fig. 2. The training involves an iterative process in which the LR weights are learned. Segmentation of a test image (light red background) is performed by its mapping to a feature space followed by a straight forward calculation of the MAP segmentation by the atlas's LR functions (Softmax regression).

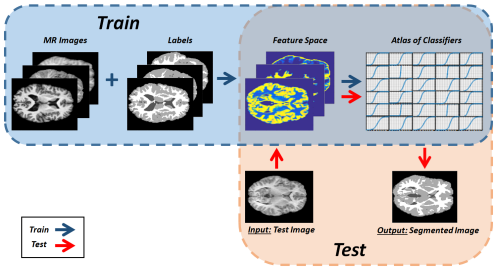

Fig. 1. An overview on the Atlas of Classifier method: In the training phase (light blue background) the MR images are projected into the feature space and are used, along with the associated label maps, to train voxel-wise classifiers. In the test phase (light red background), the test image is projected to a feature space and segmented using the trained classifiers.

2.2 Multi-Class Segmentation

Softmax regression (or multinomial LR) is a multi-class (MC) classifier [2]. Consider a set of labeled training data associated with a voxel $\mathbf{x}$, i.e. $\{\{\mathcal{I}_n(\mathbf{x}), \mathcal{Y}_n(\mathbf{x})\}\}_{n=1}^N$ where $\mathcal{I}_n(\mathbf{x}) = \{f^m(I_n(\mathbf{x}))\}_{m=1}^M \triangleq f_{\mathbf{x}n}$. This training data, can be compactly represented by a set of $(M+1)$-dimensional vectors $\mathbf{w}_{\mathbf{x}}^{\text{MC}} = \{\mathbf{w}_{\mathbf{x}}^l\}_{l=1}^L$, where $\mathbf{w}_{\mathbf{x}}^l = \{w_{\mathbf{x}}^{l,0}, w_{\mathbf{x}}^{l,1}, \ldots, w_{\mathbf{x}}^{l,M}\}$ and $w_{\mathbf{x}}^{l,0}$ is a bias parameter.

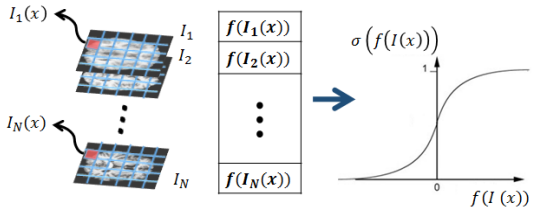

Fig. 2. Training of a classifier. Given a set of labeled images $I_1, \ldots, I_N$, a classifier associated with a voxel $\mathbf{x}$ is trained based on functions of the image intensities at that voxel $\{f(I_n(\mathbf{x}))\}_{n=1}^{N}$ and the respective labels.

Given the scalar product $\mathbf{w}_{\mathbf{x}}^{l} f_{\mathbf{x}} \triangleq w_{\mathbf{x}}^{l,0} + w_{\mathbf{x}}^{l,1} f_{\mathbf{x}}^{1} + \ldots + w_{\mathbf{x}}^{l,M} f_{\mathbf{x}}^{M}$, the multinomial LR is defined as follows:

$$p(\mathcal{Y}_*(\mathbf{x}) = y_l | I_*) = h_l(\mathbf{w}_{\mathbf{x}}^{\mathrm{MC}}, f_{\mathbf{x}}) = \frac{\exp(\mathbf{w}_{\mathbf{x}}^{l} f_{\mathbf{x}})}{\sum_{j=1}^{L} \exp(\mathbf{w}_{\mathbf{x}}^{j} f_{\mathbf{x}})} \quad \text{for } l \in \{1, \ldots, L\} \quad (2)$$

where, $\mathcal{Y}_*(\mathbf{x}) \in \{y_1, \ldots, y_L\}$, is a multi-class labeling of a voxel $\mathbf{x}$ in the given input image I_*. Thus, voxel-wise appilcation of the multinomial LRs provides soft segmentations of the examined ROIs. A hard segmentation is obtained by assigning each voxel to the class with maximum probability. The weights $\mathbf{w}_{\mathbf{x}}^{\mathrm{MC}}$ are obtained by maximizing the following objective function:

$$J(\mathbf{w}_{\mathbf{x}}^{\mathrm{MC}}) = \frac{1}{N} \sum_{n} \log p(\mathcal{Y}_{\mathbf{x}n} | f_{\mathbf{x}n}, \mathbf{w}_{\mathbf{x}}^{\mathrm{MC}}) - \frac{\lambda_R}{2} \parallel \mathbf{w}_{\mathbf{x}}^{\mathrm{MC}} \parallel^2, \quad (3)$$

where λ_R is an hyperparameter that weights the L_2 regularization term. $J(\mathbf{w}_{\mathbf{x}}^{\mathrm{MC}})$ is concave thus using gradient ascent algorithm we are guaranteed to converge to the global optimum. The LR weights update, of each class $l = 1, \ldots L$, for each $\mathbf{x} \in \Omega$, is determined by the following gradient ascent expression:

$$\mathbf{w}_{\mathbf{x}}^{l} \leftarrow \mathbf{w}_{\mathbf{x}}^{l} + \alpha \left(\frac{1}{N} \sum_{n} \left(1_{\{\mathcal{Y}_{\mathbf{x}n} = l\}} - h_l(\mathbf{w}_{\mathbf{x}}^{\mathrm{MC}}, \boldsymbol{f}_{\mathbf{x}n}) \right) \boldsymbol{f}_{\mathbf{x}n} - \lambda_R \mathbf{w}_{\mathbf{x}}^{l} \right), \quad (4)$$

where α defines the learning rate and $1_{\{\mathcal{Y}_{\mathbf{x}n} = l\}} = 1$ when $\mathcal{Y}_{\mathbf{x}n} = l$ and 0 otherwise.

2.3 Features

Intensity distribution of each image I_n is modeled by a mixture of Gaussians (GMM) with parameters $\theta_n = \{\kappa_{i,n}, \mu_{i,n}, \sigma_{i,n}\}_{i=1}^{M}$, $M = 3$, where each Gaussian represents a different tissue: CSF, WM and GM. The feature space of an image I_n is determined by its GMM such that each of its voxels $\mathbf{x}$ is represented by a feature vector $\{\mathcal{L}(\kappa_{i,n}, \mu_{i,n}, \sigma_{i,n} | I_n(\mathbf{x}))\}_{i=1}^{M}$. The GMM is learned using a maximum a-posteriori version of the expectation maximization algorithm (MAP-EM) [4] that incorporates prior knowledge (obtained by utilizing the training label maps) about the intensity distribution of each tissue (across the images), to guide the EM algorithm. This representation significantly improves the performance of our classifiers and the accuracy of the preceding registration process (to be detailed next) as compared to using the original intensities as features. It also allows multi-modal, cross-modality segmentation, see Table 2.

2.4 Registration

To form the AoC all training images and the associated annotations should be aligned to a common space. This common space is chosen to be the domain of an image $I_{\hat{n}}$, that its label map $\mathcal{Y}_{\hat{n}}$ is the one with the highest sum of pairwise Dice Similarity Coefficients (DSC) with respect to all other training labels. Let $\mathcal{I}_{\hat{n}} = \{f^m(I_{\hat{n}})\}_{m=1}^{M}$ denote the vectorized representation in the feature space of $I_{\hat{n}}$. The transformation $R_{k\to\hat{n}}$ of the source image coordinates I_k to the target image domain can be calculated by solving the following minimization problem:

$$\hat{R}_{k\to\hat{n}} = \underset{R_{k\to\hat{n}}}{\arg\ \min} D_{\text{SIM}}(\mathcal{I}_{\hat{n}}, R_{k\to\hat{n}} \circ \mathcal{I}_k)), \tag{5}$$

where D_{SIM} is the sum of squared differences between the target and the source image features. In practice, for the AoC construction, rather than calculating a multi-channel similarity measure, we perform $M = 3$ registration processes for each annotated image by a pair-wise comparison of each feature separately: $D_{\text{SIM}}(f^m(I_{\hat{n}}), R_{k\to\hat{n}} \circ f^m(I_k))$. This allows to augment the data since three pairs of feature and label maps are obtained for every training image. In addition, the aligned label maps are no longer discrete, therefore $1_{\{\mathcal{Y}_{\mathbf{x}n}=l\}}$ in Eq. (4) is replaced by a 'soft label' $\tilde{\mathcal{Y}}^l_{\mathbf{x},n} \in [0, 1]$. In the test phase, the voxel-wise multinomial LR weights $\mathbf{w}^{\text{MC}}_{\mathbf{x}}$ are registered, by a non-parametric transformation $R_{\hat{n}\to\text{TEST}}$, to the image domain of the test image [13]. This ensures that the test image features are not distorted throughout the registration process. Moreover, registration of the atlas weights $\mathbf{w}^{\text{MC}}_{\mathbf{x}}$ serves as an additional regularization, as the weights are spatially smoothed due to interpolation.

3 Experimental Results

We evaluated the AoC on the IBSR18 and IBSR20 data sets from the Internet Brain Segmentation Repository (see http://www.nitrc.org and [11]). To demonstrate multi-modal and cross-modality segmentation, we used the MRBrainS13 challenge dataset [8]. In all cases, the AoC was generated in a leave-one-out (LOO) manner excluding the test image from the atlas construction.

Unimodal experiments: Both IBSR18 (18 scans, 256 ×256 ×128) and IBSR20 (20 scans, 256 ×256 ×60) datasets consist of MRI T1-weighted volumetric images of normal subjects and manual labels for three tissue types (CSF, GM, WM). IBSR18 dataset is provided following skull stripping and bias correction via CMA routines. We used the SPM8 package (http://www.fil.ion.ucl.ac.uk/spm) for bias correction of the IBSR20. Image intensities were normalized to improve contrast between tissues. Fig. 3 visually demonstrates the process of training the multinomial LR weights, for the bias and the GM, WM intensity features. Note that as the training proceeds, the spatial contrast between the weights representing different tissues is enhanced. Also note that voxels at different anatomical locations that belong to the same tissues, may have different weight combinations. This motivates the construction of a spatial map of classifiers rather than using a single classifier for each tissue. We compared our AoC method to four other segmentation methods: GMM-EM [5], MAP-EM [4], the GMM-based

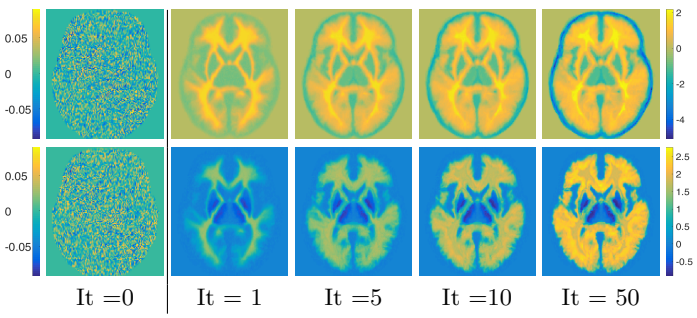

Fig. 3. Spatial maps of three Softmax weights $\mathbf{w}_{\mathbf{x}}^{l}$ of the WM class: the bias (first row) and WM (second row) intensity features, as captured during the training process based on IBSR18 dataset after 0, 1, 5,10 and 50 iterations. Left colorbars refer to the left-most images, right colorbars to the other images in the row.

Data set	Type	GMM	MAP-EM	SPM8	FSL	AoC
IBSR18	WM	77.61±6.73	85.89±2.90	89.03±1.12	87.12±3.04	**90.87±2.02**
	GM	86.27±2.98	88.46±3.01	85.19±3.01	78.59±2.50	**92.37±2.09**
	CSF	69.06±8.50	72.34±5.04	59.28±6.60	53.31±5.53	**75.62±4.19**
IBSR20	WM	68.04±14.68	76.83±11.86	81.11 ±6.29	80.10±7.54	**84.35±4.72**
	GM	77.30±18.64	84.79±5.13	88.92 ±3.69	69.64±6.38	**90.94±2.67**
	CSF	28.47±11.05	36.52±14.67	55.99±17.92	14.07±4.88	**79.48±4.23**

Table 1. DSC results for IBSR18 and IBSR20 data sets (mean ± standard deviation).

unified registration-segmentation algorithm from the SPM8 package[1] and the FAST algorithm from the FSL package [18]. The intensity distribution parameters of the training images were obtained using the ground-truth labels. The calculated parameters were used to initialize the GMM-EM and as priors for the MAP-EM of the test images. SPM8 toolbox was used with probabilistic atlases calculated from the IBSR datasets, for fair comparison. Segmentation accuracy was calculated per tissue using the DSC. Training run-time for IBSR datasets (after registration) is approximately 10 minutes ($50-70$ iterations), using unoptimized Matlab code. Segmentation of a new image takes 4 seconds on average, excluding the registration of the atlas to the feature domain. Table 1 summarizes the performance of the different methods. The proposed AoC method outperforms the other examined algorithms for both data sets, presenting the highest mean DSC measures and the lowest standard deviations, Refer also to [12] for the DSC measures obtained by other methods applied to these IBSR datasets, *without* annotation correction.

We also tested the AoC for the segmentation of the IBSR20 images while performing the training with the IBSR18 dataset and obtained the following DSC measures: 82.90±5.50 for WM; 88.23±4.08 for GM and 55.62±14.79 for CSF. Although the results are not optimal for this dataset, the method still

performs better than some of the other tested methods, demonstrating one of its main benefit – the usability of a single-shot training of one dataset for the segmentation of different datasets.

Multi-modal, cross-modality experiments: We tested the proposed AoC on the training MRBrainS13 challenge data set [8], which contains five sets of multi-modal MRI scans: T1, T1-IR and T2-FLAIR. The AoC was trained (LOO) on bi-modal scans and was tested on unimodal, including cross-modality scans, see Table 2. We find the results very promising given the extremely small training and the fact that both T1-IR and T2-FLAIR imaging (unlike T2) drastically changes the intensity distribution of the tissues, affecting the compatibility between the train and the test feature space.

	Type	**T1 and T1 IR**	**T1 and T2 FLAIR**	**T1 IR and T2 FLAIR**
T1	WM	83.65±3.80	83.83±3.75	80.93±3.78
	GM	78.56±2.85	80.66±3.28	72.54±2.68
	CSF	83.91±1.60	85.59±1.69	79.28±2.26
T1 IR	WM	79.68±4.28	79.96±4.28	77.96±4.14
	GM	71.35±2.08	74.69±2.93	69.11±2.26
	CSF	79.48±2.37	81.57±1.78	78.15±2.64
T2 FLAIR	WM	62.85±1.86	63.55±1.60	65.00±0.96
	GM	67.93±2.29	66.73±1.92	68.38±2.27
	CSF	81.94±2.12	81.71±2.09	80.83±2.13

Table 2. DSC results (mean ± standard deviation) for MRBrainS13 challenge training data. The columns represent the two modalities used for training. Rows represent modalities used for the test.

4 Discussion

A novel approach for brain MRI segmentation, which replaces the "traditional" probabilistic atlas with an *atlas of classifiers* is presented. The atlas of classifiers, is a spatial map, defined on the image domain, in which each voxel is represented by a small set of weights, learned from a much larger set of annotated scans, that defines a multinomial LR function. Each of these voxel-wise Softmax functions is more informative than the spatial probability of that voxel, obtained by averaging the training label maps, regardless of the intensity images. On the other hand, unlike existing voxel-wise classification approaches, the spatial location of each voxel does matter. Consider for example the LR weight maps, displayed in Fig. 3. Voxels of the same tissues may have different weight combinations depending on the anatomy. In contrast to the atlas fusion framework, the AoC construction is completely independent of the test images. In fact, we demonstrate comparable segmentation results for test images acquired at significantly different settings with respect to the training.

Mapping the image intensities into a feature space based on their probability distributions allows to combine the benefits of machine learning with the

prior knowledge on the underlying anatomy and imaging. The *atlas of classifiers* can be viewed as a *light weight* neural network or a grid of artificial neurons. Its structure allows inputs as big as whole images, gaining the full contextual spatial information, which cannot be preserved using image patches. Moreover, the excellent performances obtained with a very modest training procedure- a single batch with less than twenty annotated images, a relatively small set of parameters and no back-propagation, manifest its great advantage.

Acknowledgments

This study was partially supported by the Israel Science Foundation (1638/16 T.R.R) and IDF Medical Corps (T.R.R.).

References

1. Ashburner, J., Friston, K.J.: Unified segmentation. Neuroimage 26(3), 839–851 (2005)
2. Bishop, C.M.: Pattern recognition. Machine Learning 128 (2006)
3. Fischl, et al.: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33(3), 341–355 (2002)
4. Goldberger, J., Greenspan, H.: Context-based segmentation of image sequences. TPAMI 28(3), 463–468 (2006)
5. Gupta, L., Sortrakul, T.: A gaussian-mixture-based image segmentation algorithm. Pattern Recognition 31(3), 315–325 (1998)
6. Iglesias, J.E., Sabuncu, M.R.: Multi-atlas segmentation of biomedical images: a survey. MEDIA 24(1), 205–219 (2015)
7. Leemput, V., et al.: Automated model-based tissue classification of MR images of the brain. TMI 18(10), 897–908 (1999)
8. Mendrik, et al.: MRBrainS challenge: Online evaluation framework for brain image segmentation in 3T MRI scans. Computational intelligence and neuroscience (2015)
9. Moeskops, et al.: Automatic segmentation of MR brain images with a convolutional neural network. TMI 35(5), 1252–1261 (2016)
10. Pohl, et al.: A hierarchical algorithm for MR brain image parcellation. TMI 26(9), 1201–1212 (2007)
11. Rohlfing, T.: Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. TMI 31(2), 153–163 (2012)
12. Valverde, et al.: Comparison of 10 brain tissue segmentation methods using revisited IBSR annotations. J. of Magn. Res. Imaging 41(1), 93–101 (2015)
13. Vercauteren, et al.: Diffeomorphic demons: Efficient non-parametric image registration. NeuroImage 45 (2009)
14. Wachinger, C., Golland, P.: Spectral label fusion. In: MICCAI. pp. 410–417 (2012)
15. Wang, other: LINKS: Learning-based multi-source integration framework for segmentation of infant brain images. NeuroImage 108, 160–172 (2015)
16. Wells, et al.: Adaptive segmentation of MRI data. TMI 15(4), 429–442 (1996)
17. Yaqub, et al.: Investigation of the role of feature selection and weighted voting in random forests for 3-d volumetric segmentation. IEEE transactions on medical imaging 33(2), 258–271 (2014)
18. Zhang, et al.: Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. TMI 20(1), 45–57 (2001)
19. Zhang, et al.: Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. NeuroImage 108, 214–224 (2015)