Published in final edited form as:

Mach Learn Med Imaging. 2016 October ; 10019: 237–245. doi:10.1007/978-3-319-47157-0_29.

Regression Guided Deformable Models for Segmentation of Multiple Brain ROIs

Zhengwang Wu, Sang Hyun Park, Yanrong Guo, Yaozong Gao, and Dinggang Shen

Department of Radiology and BRIC, UNC at Chapel Hill, Chapel Hill, NC, USA

Abstract

This paper proposes a novel method of using regression-guided deformable models for brain regions of interest (ROIs) segmentation. Different from conventional deformable segmentation, which often deforms shape model locally and thus sensitive to initialization, we propose to learn a regressor to explicitly guide the shape deformation, thus eventually improves the performance of ROI segmentation. The regressor is learned via two steps, (1) a joint classification and regression random forest (CRRF) and (2) an auto-context model. The CRRF predicts each voxel's deformation to the nearest point on the ROI boundary as well as each voxel's class label (e.g., ROI *versus* background). The auto-context model further refines all voxel's deformations (i.e., deformation field) and class labels (i.e., label maps) by considering the neighboring structures. Compared to the conventional random forest regressor, the proposed regressor provides more accurate deformation field estimation and thus more robust in guiding deformation of the shape model. Validated in segmentation of 14 midbrain ROIs from the IXI dataset, our method outperforms the state-of-art multi-atlas label fusion and classification methods, and also significantly reduces the computation cost.

1 Introduction

Segmenting the brain into distinct ROIs is the foundation work for establishing relationships (functional network construction) between different brain ROIs. Recently, more and more discoveries reveal the importance of analyzing brain functional networks for understanding human brain, emotion and behaviors [1]. Therefore, better brain ROI segmentation is highly demanded.

Previously, many automatic methods have been proposed for brain ROI segmentation. One popular approach is the *multi-atlas label fusion based segmentation* [2]. In this approach, atlases with pre-segmented ROI labels are registered onto the target image. The aligned labels are then fused to derive the final segmentation of target image. However, it suffers from high-computational cost, due to nonlinear registration and local correspondence searching. To compromise for the computational cost, the number of used atlases is generally small, e.g., 20 atlases. For example, this approach often selects a subset of the most relevant atlases, which unfortunately loses some potentially useful information in other unselected atlases. On the other hand, the *classification-based approach* has also been

proposed. Without relying on nonlinear registration, this approach is very efficient. However, its performance is limited since it overlooks the spatial information of brain ROIs. Moreover, classification based approach may produce weird ROI shapes and unsmooth boundaries, since the labeling of neighboring voxels is independent and no shape constraint is imposed. To address this limitation, Ma et al. [3] combined these two approaches into a single framework by extracting additional features from aligned atlases to train ROI classifiers. However, similar to the multi-atlas based approach, the efficiency is still a big concern due to the need of nonlinear registration.

In this paper, we propose a novel deformable model to segment brain ROIs. Different from conventional methods, which search local boundaries to deform the shape model, we explicitly learn a regressor to guide the shape deformation inspired by, which firstly addressed the regressor guided deformable segmentation. The difference between our work and [4] is that we introduced the context features to refine learned regressor. Specifically, we propose a joint classification and regression random forest (CRRF) regressor, which predicts both the deformation and ROI label of each voxel. Experimental results show that the incorporation of ROI classification into deformation estimation framework could significantly improve the deformation estimation in CRRF. To further impose the structured information during voxel-wise deformation prediction, an auto-context model is also adopted. It iteratively refines the estimated deformation fields by employing the neighboring predictions as new features to refine the training of the regressor. Compared with existing methods, our method has three advantages: (1) it is registration-free and thus very efficient; (2) it does not suffer from initialization problem as in the conventional deformable models; (3) it can impose both global-shape and local-smoothness constraints during the deformable ROI segmentation. We applied our method to IXI data set for segmenting 14 midbrain ROIs. The quantitative comparison between our results and manual segmentations proved the effectiveness, and also demonstrated its outstanding performance over several state-of-the-art methods.

2 Method

2.1 Main Framework

The main framework of our method is illustrated in Fig. 1. For each ROI, a shape model (represented as a triangle mesh in the 3D space [5]) is initialized to the image center and then iteratively deformed to the target ROI boundary under the guidance of the estimated deformation field (DF). Here, the DF is efficiently estimated by a CRRF and refined in the auto-context model, which will be detailed in Sect. 2.2. Section 2.3 presents how the estimated DF guides the shape model's deformation to the target ROI iteratively. In this paper, a multi-resolution strategy is also adopted to increase both the robustness and efficiency of our proposed method. Specifically, the coarse-level DF is first estimated in the down-sampled image space. It is used to roughly guide the shape model onto the ROI boundary. Then, the fine-level DF is estimated only in a local region surrounding the ROI for voxel-level refinement, thus improving both the efficiency of segmentation in the fine resolution, as well as the robustness of segmentation by using the coarse-level DF to guide the deformation.

Based on the above framework, each brain ROI can be independently segmented by the proposed deformable model. This can be done in a parallel way, which could greatly improve the segmentation efficiency. However, it potentially brings the issue of overlapping among different ROIs. To address this issue, a merging step is further adopted to combine all segmented ROIs into a single multi-label segmentation, which is detailed in Sect. 2.4.

2.2 Deformation Field Estimation

This section is organized as follows. First, we summarize overall pipeline of deformation field estimation, including the definition of deformation at each voxel and also the sampling strategy used in the training stage. Then, we present a joint classification and regression random forest (CRRF) for estimating both the deformation and ROI label of each voxel. Finally, we elaborate how to combine CRRF with auto-context for iteratively refining the deformation field.

Overall Pipeline—The deformation at each voxel is defined as a 3D displacement from this voxel to the nearest voxel on the ROI boundary, as illustrated in Fig. 2. It is given in the training images, but unknown for a testing image. In the *training stage*, we uniformly sample a set of voxels from training images. For each voxel, we compute its deformation and extract its ROI label, and then use them as the ground-truth prediction targets. The appearance features for each voxel are the 3D Haar-like features [6] extracted from the local patch. The choice of Haar-like features is due to its computational efficiency by using the integral image. Given features and prediction targets, we can use them to train a CRRF for each brain ROI. In the *testing stage*, each voxel of the testing image will be visited. Specifically, the appearance features of each visited voxel are extracted from the local patch and then fed into the learned CRRF to estimate the deformation at this voxel towards a brain ROI. Through this voxel-wise estimation, a deformable field can be finally obtained for each brain ROI under segmentation. To iteratively refine the deformation field by auto-context, additional 3D Haar features are extracted from the intermediate deformation field. Since they encode the neighboring prediction information, the combination of these features with original appearance features could lead to the structured refinement of the deformation field. This refinement step can be iteratively conducted for updating the deformation field.

Sampling/Training Strategy—Figure 3 illustrates the sampling strategies used in the low and high resolutions. In the low resolution, the training voxels are uniformly sampled all over the whole image, and a single CRRF is trained to jointly predict the deformations of a voxel to all ROIs. In the high resolution, the training voxels are randomly sampled near each ROI, and CRRF is trained separately for each ROI. The joint CRRF training can encode the spatial relationship between ROIs, thus improving the robustness of localization in the low resolution. In the high resolution, the specificity of the CRRF can be improved by sampling around the individual ROIs for training ROI-specific CRRFs.

Joint Classification and Regression Random Forest (CRRF)—To adapt random forest for joint classification and regression, the objective function of random forest at each node is modified as follows:

$$\arg \max_{f, \tau} \frac{1}{Z_V} (V - \sum_{q \in \{L, R\}} \frac{N^q}{N} V^q) + \frac{1}{Z_E} (E - \sum_{q \in \{L, R\}} \frac{N^q}{N} E^q)$$

where f and τ are the pair of feature and threshold of one node to be optimized. V , E , N denote the average variance, entropy and the number of training samples at this node. The superscript $q \in \{L, R\}$ indicates the measurement computed after splitting this node into left(L) and right(R) child nodes. Z_V and Z_E are the average variance and entropy at the root node, respectively. The first term in the above formula is for the deformation regression, and the second term is for the ROI classification. With the above modification, a single random forest can be utilized to simultaneously predict the deformation and ROI label of a voxel. Similar to the theory of multi-task learning, we find that the joint classification and regression random forest produces better deformation field than the conventional regression forest, as illustrated in Fig. 4.

Iterative Auto-Context Refinement—With CRRF, the deformation at each voxel can be predicted from the local appearance. However, independent voxel-wise estimation neglects the relationship among deformations of neighboring voxels. Thus, the estimated deformation field is often noisy and unsmooth. To overcome it, an auto-context model [7] is adopted to iteratively refine the deformation field. In auto-context model, a sequence of CRRF is trained iteratively, as schematically illustrated in Fig. 5. In iteration **0**, the CRRF is trained with only appearance features. Once the first CRRF is obtained, it can be used to estimate a tentative deformation field as well as a ROI probability map. In iteration **1**, additional 3D Haar features are extracted from both the tentative deformation field and ROI probability map. These features are context features, because they encode both the neighboring deformation and ROI label information. By combining context features with appearance features, a second CRRF can be trained to refine both the deformation field and ROI probability map. This procedure can be repeatedly conducted until the maximum iteration T is reached.

2.3 Regression Guided Deformable Segmentation

The mean ROI shape model is first initialized at the image center. Then, based on the estimated deformation field $\mathbf{D}$, we propose a hierarchical strategy to deform the shape model under the guidance of the estimated deformable field. Denote the vertex set of the shape model as $\mathcal{V} = \{\mathbf{p}_1, \dots, \mathbf{p}_m, \dots, \mathbf{p}_{|\mathcal{V}|}\}$, where $\mathbf{p}_m$ is the m -th vertex and $|\mathcal{V}|$ is the total number of vertices. Let $\mathbf{D}_{\mathbf{p}_m}$ denote the estimated deformation for vertex $\mathbf{p}_m$ in the image domain. The whole deform procedure is conducted in the following 3 steps, *First*, a translation vector

is estimated as the average deformation of all vertices, i.e., $\Delta = \sum_{m=1}^{|\mathcal{V}|} \mathbf{D}_{\mathbf{p}_m} / |\mathcal{V}|$, which brings the deformable model into the correct position of ROI. *Second*, a rotation matrix $\mathbf{R}$ is estimated between the deformed shape model $\hat{\mathcal{V}}$ and the original shape model $\mathcal{V}$, which

brings the deformable model to the correct orientation, i.e., $\arg \min_{\mathbf{R}} \|\hat{\mathcal{V}} - \mathbf{R}\mathcal{V}\|_2^2$ s.t., $\mathbf{R}\mathbf{R}^T = \mathbf{I}$, where $\hat{\mathcal{V}} = \{\mathbf{p}_1 + \rho \mathbf{D}_{\mathbf{p}_1}, \dots, \mathbf{p}_m + \rho \mathbf{D}_{\mathbf{p}_m}, \dots, \mathbf{p}_{|\mathcal{V}|} + \rho \mathbf{D}_{\mathbf{p}_{|\mathcal{V}|}}\}$, and ρ is the step size.

Similarly, an affine transformation matrix can be also estimated to correct the scaling and

shearing parameters of the deformable model. *Finally*, once the pose of deformable model is well estimated, the deformable model can be freely deformed locally under the deformation field. To maintain the mesh quality, the deformable model is further smoothed and re-meshed after each iteration of freeform deformation. With an explicit deformation field, the proposed deformable model is able to improve both the robustness to initialization and the flexibility of deformation.

2.4 ROI Merging

Once all ROIs are independently segmented by the proposed regression-guided deformable models, the final step is to address the potential overlapping issue by fusing all the individual ROI segmentations together. To merge all ROI segmentations, we utilize the ROI probability maps from CRRF. Note that each ROI probability map tells the probability of a voxel belonging to the ROI. Thus, for a voxel in the overlapping region, its ROI label can be determined by the one with the maximum probability value across all ROIs.

3 Experiments

Our method is evaluated in the IXI dataset [8] with 30 subjects. In this study, 14 midbrain ROIs are used for evaluating the segmentation performance, i.e., Right(**R**)/Left(**L**) hippocampus, **R/L** amygdala, **R/L** caudate, **R/L** nucleus(nuc.) accumbens, **R/L** putamen, **R/L** thalamus and **R/L** globus pallidus. The 3-fold cross validation is used to evaluate the performance. At each fold, 20 subjects are used for training and 10 subjects are used for testing. The quantitative assessment is presented by calculating the Dice ratio between the automatic segmentations and manual ROI segmentations. Five state-of-the-art segmentation methods are compared with our method, including, (a) **m**ajority **v**oting (MV) [9]; (b) **l**ocal **w**eighted **v**oting (LWV) [10]; (c) **p**atch based **l**abel **f**usion (PLF) [11]; (d) **s**pase **p**atch based **l**abel fusion (SPLF) [12]; (e) **r**andom **f**orest **c**lassification (RFC) [3].

All comparison results are reported in Table 1. It can be seen that our method consistently achieves the best performance for all 14 ROIs (with the mean Dice ratio of 0.83). Compared to the results of multi-atlas based methods (columns 2–5), our method has improved the Dice ratio by 0.07 on average. This indicates the effectiveness of our method. Furthermore, the computational time of our method (about 3 min) for segmenting all 14 ROIs is also much less than those of multi-atlas based methods (about 10–15 minutes on average, with the same number of atlases as used as training images in our method). Column 6 shows the segmentation results using random forest classification, with both the same multi-resolution training strategies and the same auto-context refinement. Although this simple random forest classification method can run as fast as our method, its performance is generally limited because it overlooks the spatial information of ROIs and also does not enforce the shape constraints of the ROIs during the segmentation. Our method obtains higher average dice ratio by 0.08 than the simple random forest classification.

Typical segmentation result using our proposed method is also demonstrated in Fig. 6. It can be seen, compared to the manual labels, our results are smoother. This is understandable, since manual ROI labeling need to validate at 3 cross-sectional views during manual delineation, the smoothness is thus difficult to maintain. For the RFC method, since each

voxel is independently segmented, there are some mis-classifications, as shown in Fig. 6. Clearly, the ROIs segmented by our method are more coherent and do not contain the isolated voxels.

4 Conclusion

In this paper, we have proposed learning a regressor to explicitly predict the deformation field for guiding the deformable segmentation of multiple brain ROIs. Specifically, a joint classification and regression random forest (CRRF) is developed to estimate both the deformation and label of each voxel. Compared to the conventional regression forest, CRRF provides better deformation estimations. To tackle the limitation of independent voxel-wise estimation, an auto-context model is further adopted to iteratively refine the estimated deformation field by considering the neighboring prediction results. Validated on the 14 midbrain ROIs of IXI dataset, our method outperforms other five multi-atlas based label fusion and classification approaches.

Acknowledgments

D. Shen—This work was supported by the National Institute of Health grants 1R01 EB006733.

References

1. Satterthwaite TD, Davatzikos C. Towards an individualized delineation of functional neuroanatomy. *Neuron*. 2015; 87:471–473. [PubMed: 26247857]
2. Aljabar P, Heckemann RA, et al. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage*. 2009; 46:726–738. [PubMed: 19245840]
3. Ma, G., Gao, Y., Wu, G., Wu, L., Shen, D. Atlas-guided multi-channel forest learning for human brain labeling. In: Menze, B.Langs, G.Montillo, A.Kelm, M.Müller, H.Zhang, S.Cai, WT., Metaxas, D., editors. MCV 2014. LNCS. Vol. 8848. Springer; Heidelberg: 2014. p. 97-104. Revised Selected Papers
4. Glocker, B., Pauly, O., Konukoglu, E., Criminisi, A. Joint classification-regression forests for spatially structured multi-object segmentation. In: Fitzgibbon, A.Lazebnik, S.Perona, P.Sato, Y., Schmid, C., editors. ECCV 2012, Part IV. LNCS. Vol. 7575. Springer; Heidelberg: 2012. p. 870-881.
5. Cootes TF, Taylor CJ, et al. Active shape models-their training and application. *Comput Vis Image Underst*. 1995; 61:38–59.
6. Cui, X., Liu, Yea. ICME-2007. IEEE; 2007. 3D HAAR-like features for pedestrian detection; p. 1263-1266.
7. Tu Z, Bai X. Auto-context and its application to high-level vision tasks and 3D brain image segmentation. *IEEE T-PAMI*. 2010; 32:1744–1757.
8. Brain Development Org. IXIDataset. <http://www.brain-development.org>
9. Wang H, Suh JW, Das SR, Pluta JB, Craige C, Yushkevich PA. Multi-atlas segmentation with joint label fusion. *IEEE T-PAMI*. 2013; 35:611–623.
10. Artaechevarria X, Munoz A, et al. Combination strategies in multi-atlas image segmentation: application to brain MR data. *IEEE T-MI*. 2009; 28:1266–1277.
11. Coupé P, Manjón JV, et al. Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. *NeuroImage*. 2011; 54:940–954. [PubMed: 20851199]
12. Wu G, Wang Q, Zhang D, Shen D. Robust patch-based multi-atlas labeling by joint sparsity regularization. *MICCAI Workshop STMI*. 2012

Fig. 1. Main framework of our proposed method. DF denotes deformable field. **LR** and **HR** denote low resolution and high resolution, respectively.

Fig. 2.
Deformation definition.

Fig. 3.
Two sampling strategies.

Fig. 4.

Qualitative comparison between deformation fields (DFs) estimated by conventional regression forest (middle) and our CRRF (right).

Fig. 5.
Illustration of iterative auto-context refinement.

Fig. 6.
Qualitative segmentation comparison obtained by RFC and our method.

Table 1

Comparison with other methods. Bolded numbers indicate the best performance.

Method	MV	LWV	PLF	SPLF	RFC	Ours
R. Hippocampus	0.71	0.76	0.81	0.80	0.70	0.83
L. Hippocampus	0.70	0.73	0.78	0.79	0.72	0.81
R. Amygdala	0.71	0.74	0.77	0.75	0.71	0.78
L. Amygdala	0.73	0.73	0.79	0.76	0.73	0.80
R. Caudate	0.80	0.82	0.86	0.83	0.73	0.89
L. Caudate	0.80	0.83	0.86	0.83	0.82	0.90
R. nuc. accumbens	0.58	0.60	0.64	0.68	0.63	0.72
L. nuc. accumbens	0.57	0.60	0.64	0.68	0.60	0.71
R. Putamen	0.81	0.83	0.86	0.83	0.83	0.88
L. Putamen	0.80	0.83	0.86	0.83	0.79	0.88
R. Thalamus	0.84	0.85	0.88	0.84	0.88	0.91
L. Thalamus	0.84	0.86	0.88	0.84	0.85	0.90
R. Globus pallidus	0.76	0.77	0.80	0.79	0.77	0.82
L. Globus pallidus	0.72	0.75	0.78	0.78	0.72	0.83