Brain Image Labeling Using Multi-atlas Guided 3D Fully Convolutional Networks

Longwei Fang[1,2,4], Lichi Zhang[4], Dong Nie[4], Xiaohuan Cao[4,5],
Khosro Bahrami[4], Huiguang He[1,2,3(✉)], and Dinggang Shen[4(✉)]

[1] Research Center for Brain-Inspired Intelligence, Institute of Automation,
Chinese Academy of Sciences, Beijing, China
huiguang.he@ia.ac.cn
[2] University of Chinese Academy of Sciences, Beijing, China
[3] Center for Excellence in Brain Science and Intelligence Technology,
Chinese Academy of Sciences, Beijing, China
[4] Department of Radiology and BRIC,
University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
dinggang_shen@med.unc.edu
[5] School of Automation, Northwestern Polytechnical University, Xi'an, China

Abstract. Automatic labeling of anatomical structures in brain images plays an important role in neuroimaging analysis. Among all methods, multi-atlas based segmentation methods are widely used, due to their robustness in propagating prior label information. However, non-linear registration is always needed, which is time-consuming. Alternatively, the patch-based methods have been proposed to relax the requirement of image registration, but the labeling is often determined independently by the target image information, without getting direct assistance from the atlases. To address these limitations, in this paper, we propose a multi-atlas guided 3D fully convolutional networks (FCN) for brain image labeling. Specifically, multi-atlas based guidance is incorporated during the network learning. Based on this, the discriminative of the FCN is boosted, which eventually contribute to accurate prediction. Experiments show that the use of multi-atlas guidance improves the brain labeling performance.

1 Introduction

Accurate labeling of neuro-anatomical regions is highly demanded for quantitative analysis of MR brain images. Many attempts have been made in automatic labeling methods since it is infeasible to manually label a large set of 3D MR images. However, it remains a challenging problem due to the complicated brain structures and also the ambiguous boundaries between some regions of interest (ROIs).

The multi-atlas based methods have emerged as the standard way in the brain image labeling for its effectiveness and robustness. By using the atlases, each with a single MRI scan and its manual label maps, the multi-atlas based methods first *register* multiple atlases to the target image and then *fuse* the respective deformed atlas label maps to obtain the labeling results. Many relevant works have been made to improve the performances of these *registration* and *label fusion* steps in the multi-atlas based

methods, as summarized in [1–3]. However, one major limitation of these multi-atlas based method is that it always needs non-rigid registration for aligning atlases to the subject, which is time-consuming [4]. Besides, it is also a challenging work to obtain accurate registration, which will eventually affect the final labeling performance.

On the other hand, the patch-based methods have gained increased attentions recently, which are mainly developed to relax the high demands of registration accuracy in the multi-atlas based methods. Specifically, in the patch-based methods, each patch in the target subject image looks for its similar patches in the atlas images according to patch similarity. Then, the label of those selected atlas patches are fused together to label the center voxel of subject patch [5, 6]. The weights of selected atlas patches in the label fusion process are estimated based on their intensity similarity with the target subject patch. Also, Wu et al. [7] further proposed using a multi-scale feature representation and label-specific patch partition method to extend the label fusion strategy. In this method, each patch is represented by the multi-scale features that encode both local and semi-local image information, and then the image patch is further partitioned into a set of label specific partial image patches. Finally, the hierarchical patch-based label fusion is followed to finish the labeling. On the other hand, the learning-based methods have also been incorporated into the brain image labeling process, generally in a patch-based manner. For example, Tu and Bai [8] extracted the 3D Haar features from the atlases and then employed the probabilistic boosting tree (PBT) to learn the classifier for brain labeling. Hao et al. [9] introduced a hippocampus segmentation method using L1-regularized support vector machine (SVM), with a k-nearest neighbor (kNN) based training sample strategy. Moreover, the random forest has also been widely applied, since it can efficiently handle a large number of training atlases, and can largely avoid the overfitting problem in the conventional decision tree methods by incorporating the uniform bagging strategy [10, 11]. Recently, fully convolutional networks (FCN) [12] have shown excellent performance in natural image segmentation and recognition. Some researchers have also employed the FCN model for medical image segmentation. For example, Nie et al. [13] adopted the FCN model for brain tissue segmentation, which has shown a promising result.

However, the main limitation of the current methods is that they determine the target labels merely on the local appearance of target image patch, without considering the direct label information from those similar atlas patches. Besides, although patch-based methods can relax the demand of accurate registration, most methods [6–10] still apply *non-rigid* registration to preprocessing the data, for the benefit of labeling improvements.

In this paper, we intend to solve the aforementioned issues by proposing a multi-atlas guided 3D FCN model for improving the performance of brain labeling. The major contribution here is two-fold. *First*, we develop a novel multi-atlas guidance strategy, which can directly utilize prior information in the atlases to guide and improve the labeling capability. *Second*, different from the conventional multi-atlas based methods, we need no *non-rigid* registration for aligning atlases to the target image, by still guaranteeing the reasonable labeling performance. This will greatly reduce the time cost for the overall labeling process, thus making it more applicable for future clinical applications.

2   Methods

In this section, we will illustrate the details of our proposed multi-atlas guided FCN method, which consists of the *training* and *testing* stages. In the *training* stage, we first select a number of images from the training set, and consider them as the atlas images. Then, we extract 3D cubic patches from the training images, and, for each selected training patch, we also select K most similar atlas patches from the linearly-aligned atlas images. Next, each training patch and its corresponding selected atlas patches (including intensity patches and label patches) are used together to train the FCN model. In the *testing* stage, the trained FCN model is first applied to each input testing patch (of the new testing image) and its selected atlas patches, for obtaining a predicted label patch. Then, all the predicted label patches from all locations of the testing image are fused together to give the final labeling result.

2.1   Training Data Preparation

Data Preprocessing: The first step is normalizing the intensity of data in the range from 0 to 255. And before the patch extraction process, for each training image, we first register all atlases to its space. As stated above, we need no *non-rigid* registration; instead, we just use affine registration, which can be implemented more efficiently. Specifically, we first linearly align the intensity images of atlases to the target training image using the *flirt* in FSL [14], and warp the label maps of all atlases to the training image space by using the obtained respective linear transformation for each atlas.

Patch Extraction: Since there are high variations of ROI sizes for different brain ROIs under labeling, we develop a specific patch extraction strategy to ensure that the sufficient training patches can be extracted from each ROI under labeling. Specifically, this strategy ensures an adequate number of patches extracted around the boundary of each ROI, since boundaries contain the direct shape information vital for ROI labeling. To do this, we first employ a canny edge detector to find boundaries in each of the atlas label maps. Then, we randomly select the patches by ensuring that (1) the number of patches extracted from every ROI is similar, and (2) the number of patches extracted from the boundary of each ROI is similar to the number of patches extracted from internal part of each ROI.

Atlas Patch Selection: For each given training image patch $P_{T(I,j)}$, centered at voxel j and extracted from the training image I, we can find one most similar atlas image patch from each atlas in the 3D cubic searching neighborhood $c(j)$, i.e., according to the image intensity similarity. This step can be mathematically summarized by Eq. 1, where (M, n) is an atlas image patch selected from the atlas image at the location of voxel n, and $|| \cdot ||_2$ is a Euclidean distance measure between image patches under comparison.
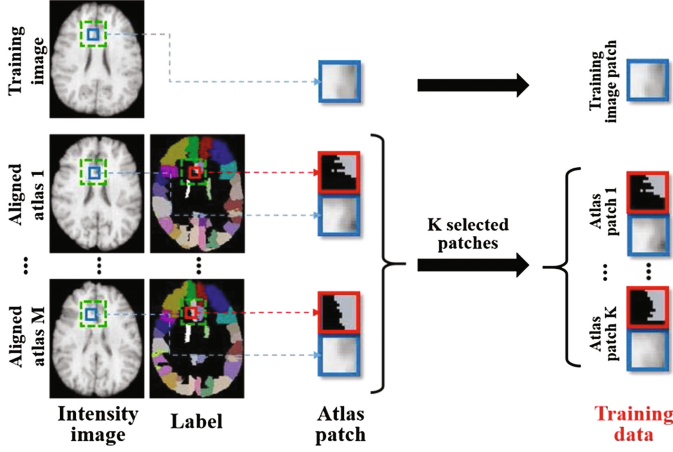
Fig. 1. A brief illustration of steps for preparing the training data. The green dash box is the searching neighborhood. (Color figure online)

$$\hat{P} = \{P_{A(M,n)} \mid \min_{n \in c(j)} ||P_{T(I,j)} - P_{A(M,n)}||_2^2\} \qquad (1)$$

By ranking all the selected atlas image patches according to their respective similarities to the training image patch $(I,)$, we can finally select the top K (i.e., $K = 3$) atlas image patches. Then, each training image patch and its K selected atlas image patches are combined as joint input to train our proposed FCN model. Figure 1 summaries all steps in our method for prepressing the training data to train the FCN model.

2.2 Fully Convolutional Networks (FCN) Configuration

We employ an FCN model for the brain ROI labeling. FCN model is an end-to-end learning structure, with its output as a patch. Compared with the convolutional neural networks (CNN) [16] that output is just the label for the center voxel of the input image patch, FCN can label the whole patch in one process, thus more efficient and potentially more spatially-consistent labeling than CNN. The configuration of our FCN (as shown in Fig. 2) is briefed below. (1) We first learn $K + 1$ mapping structures separately for the training image patch and K selected atlas image/label patches. Specifically, in the first layer, for each of K sets of selected atlas image/label patches, we use K concatenated layers to group the image patch and label patch of the same atlas together. For the training image patch, since there is no label patch, it is simply input the FCN. Next, three convolution layers are applied to each of $K + 1$ mapping structures, followed by a max pooling layer for down sampling the mapped data. (2) After separately mapping the training image patch and the K selected atlas image/label patches, we use another concatenation layer to combine $K + 1$ sets of mapped data together, followed by two
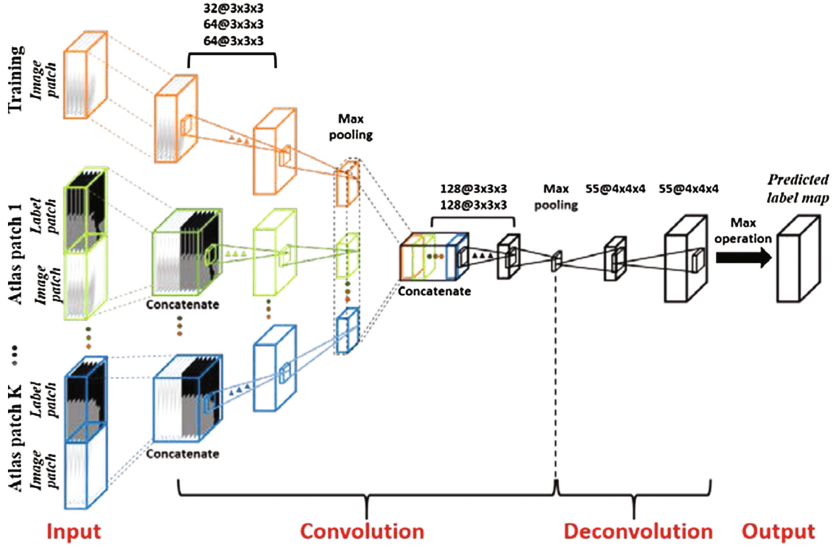
Fig. 2. Detailed structure and parameters of our proposed FCN model for patch labeling.

convolution layers and a max pooling layer. (3) Finally, we use two deconvolution layers to get the label map. Note that the rectified linear units (ReLU) is used as our activation function for all the convolution layers, and also cross-entropy loss is used as our loss function.

2.3   Brain Labeling

For each new testing brain image, we first use affine registration to align all the atlases to this target image. Then, for each (testing) image patch (with the same size as all the training image patches) extracted from the testing image, we select its K most similar atlas image patches from all linearly-aligned atlases as described in Sect. 2.1. Next, each testing image patch and its K selected atlas image/label patches are combined and inputted to our trained FCN for obtained the patch labeling result. Finally, the labeling results from all testing patches covering the whole testing image are fused together (with majority voting) to produce a final label map for the testing image.

3   Experimental Results

We use the LONI LPBA40[1] dataset to evaluate the performance of our proposed brain ROI labeling method. The LONI LPBA40 dataset contains 40 T1-weighted MR brain images with 54 manually labeled ROIs. In our method, four-fold cross validation is

---

[1] http://www.loni.ucla.edu/Atlases/LPBA40.

used. Specifically, in each fold, we select 10 images as the testing images, and the rest as the training images. Furthermore, we select 10 images from those 30 training images as the atlas images, and other 2 images as the validation images for FCN training. Note that we also train another FCN model without using multi-atlas guidance (i.e., just using the training image patch), and use it as the baseline method. Note that the network structures and parameters are same in both our proposed multi-atlas guided FCN method and this baseline method. In our paper, we use the patch size of $24 \times 24 \times 24$ in voxels, and the searching neighborhood size of $30 \times 30 \times 30$ also in voxels. The number of training image patches sampled from each training image is 8,400. For the testing image, we evenly visit patches with a step size of 9 voxels, to ensure a sufficient overlap for the neighboring patches.

We evaluate the labeling performance using the Dice Similarity Coefficient (DSC). The results on LONI LPBA40 show that our proposed method can achieve the average DSC of $(80.33 \pm 1.26)\%$ for 54 ROIs. Table 1 lists the comparison of our method with the state-of-art methods. Note that, for these state-of-art methods, we simply copied results from [7, 10, 11] for fair comparison. It can be observed that our proposed method outperforms the non-local based method [11] for more than 2%, and also achieves a comparable labeling results to the non-rigid registration methods [7, 10]. Although the mean DSC estimations by the multi-atlas method [10] and our proposed method are close, it can be observed that our method has a much smaller standard deviation, suggesting that our method is more reliable. Furthermore, it often takes 2–20 h for just the non-rigid registration step in multi-atlas method [15], while our proposed method takes less than 15 min for labeling a testing image which is definitely more efficient in the application stage.

Table 1. Quantitative comparison between the proposed method and the state-of-arts methods.

Method	Non-rigid registration		Affine registration		
	Multi-atlases [7]	Learning [10]	Non-local [11]	FCN-single patch	Proposed
DSC (%)	81.46 ± 2.25	80.1 ± 4.53	78.26 ± 4.83	78.20 ± 1.60	80.33 ± 1.26

We further compared our method with the baseline method (namely FCN-single patch) in Table 1, which shows significant improvements for ROI labeling using multi-atlas guidance in our method. The structure of baseline method is similar with proposed method, except that baseline method does not have atlas patches. Figure 3 also shows a labeled testing image by the baseline method (FCN-single patch) and our proposed method (Proposed). Figure 3(a) shows the golden standard (obtained with manual delineation). Figure 3(b) shows the labeling result by the baseline method (FCN-single patch), and Fig. 3(c) shows the labeling result by our proposed method (Proposed). It can be observed that, the labeling results on the boundary by proposed method is smoother than the baseline method. Moreover, there are wrong predictions inside of some ROIs by the baseline method, as indicated in Fig. 3(b). When using multi-atlas guidance to train the FCN model in our proposed method, more prior labeling information from multiple atlases can be used to directly help refine the labeling results, thus avoiding the wrong labeling by the baseline method.
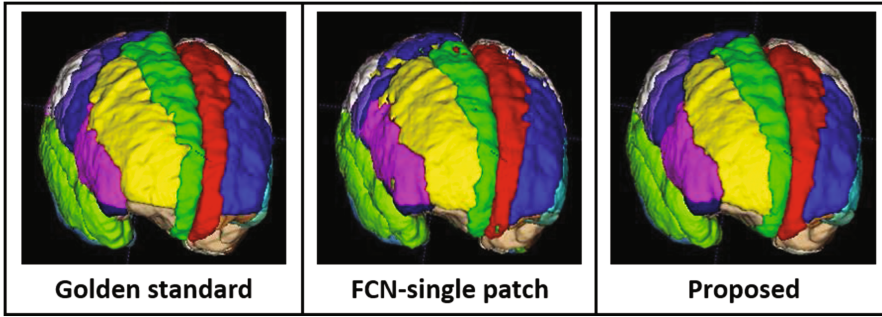
Fig. 3. Visual comparison of labeling results by the baseline method (FCN-single patch) and our proposed method (Proposed).

4   Conclusion

In this paper, we have presented a multi-atlas guided 3D FCN method for brain ROI labeling. Different from the traditional neural networks, the input to our FCN includes *not only* the intensity image patch from training (or testing) image, *but also* both the intensity and label patches from the atlases. Such combination can provide a clearer guidance for FCN to better label the target brain images. Furthermore, our proposed method requires no non-rigid registration for data preprocessing. The validation results on a public dataset show that our proposed method outperforms the non-local based methods in accuracy and non-registration based methods in speed, as well as the baseline method in terms of labeling accuracy.

References

1. Jia, H., Yap, P.T., Shen, D.: Iterative multi-atlas-based multi-image segmentation with tree-based registration. NeuroImage **59**(1), 422–430 (2012)
2. Wolz, R., Aljabar, P., Hajnal, J.V., Hammers, A., Rueckert, D., The Alzheimer's Disease Neuroimaging Initiative: LEAP: learning embeddings for atlas propagation. NeuroImage **49**(2), 1316–1325 (2010)
3. Langerak, T.R., van der Heide, U.A., Kotte, A.N., Viergever, M.A., Van Vulpen, M., Pluim, J.P.: Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE). IEEE Trans. Med. Imaging **29**(12), 2000–2008 (2010)
4. Iglesias, J.E., Sabuncu, M.R.: Multi-atlas segmentation of biomedical images: a survey. Med. Image Anal. **24**(1), 205–219 (2015)
5. Coupé, P., Manjón, J.V., Fonov, V., Pruessner, J., Robles, M., Collins, D.L.: Patchbased segmentation using expert priors: application to hippocampus and ventricle segmentation. NeuroImage **54**(2), 940–954 (2011)
6. Wang, H., Suh, J.W., Das, S.R., Pluta, J.B., Craige, C., Yushkevich, P.A.: Multiatlas segmentation with joint label fusion. IEEE Trans. Pattern Anal. Mach. Intell. **35**(3), 611–623 (2013)

7. Wu, G., Kim, M., Sanroma, G., Wang, Q., Munsell, B.C., Shen, D., The Alzheimer's Disease Neuroimaging Initiative: Hierarchical multi-atlas label fusion with multi-scale feature representation and labelspecific patch partition. NeuroImage **106**, 34–46 (2015)
8. Tu, Z., Bai, X.: Auto-context and its application to high-level vision tasks and 3D brain image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **32**(10), 1744–1757 (2010)
9. Hao, Y., Wang, T., Zhang, X., Duan, Y., Yu, C., Jiang, T., Fan, Y.: Local label learning (LLL) for subcortical structure segmentation: application to hippocampus segmentation. Hum. Brain Mapp. **35**(6), 2674–2697 (2014)
10. Zikic, D., Glocker, B., Criminisi, A.: Encoding atlases by randomized classification forests for efficient multi-atlas label propagation. Med. Image Anal. **18**(8), 1262–1273 (2014)
11. Zhang, L., Wang, Q., Gao, Y., Wu, G., Shen, D.: Automatic labeling of MR brain images by hierarchical learning of atlas forests. Med. Phys. **43**(3), 1175–1186 (2016)
12. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440 (2015)
13. Nie, D., Wang, L., Gao, Y., Sken, D.: Fully convolutional networks for multimodality isointense infant brain image segmentation. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 1342–1345. IEEE (2016)
14. Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S.M.: Bayesian analysis of neuroimaging data in FSL. Neuroimage **45**(1), S173–S186 (2009)
15. Landman, B., Warfield, S.: Miccai 2012 workshop on multi-atlas labeling. In: Medical Image Computing and Computer Assisted Intervention Conference 2012: MICCAI 2012 Grand Challenge and Workshop on Multi-Atlas Labeling Challenge Results (2012)
16. Moeskops, P., Viergever, M.A., Mendrik, A.M., de Vries, L.S., Benders, M.J., I˘sgum, I.: Automatic segmentation of MR brain images with a convolutional neural network. IEEE Trans. Med. Imaging **35**(5), 1252–1261 (2016)

Patch-Based Techniques in Medical Imaging
Third International Workshop, Patch-MI 2017, Held in
Conjunction with MICCAI 2017, Quebec City, QC, Canada,
September 14, 2017, Proceedings
Wu, G.; Munsell, B.C.; Zhan, Y.; Bai, W.; Sanroma, G.;
Coupé, P. (Eds.)
2017, XI, 168 p. 59 illus., Softcover
ISBN: 978-3-319-67433-9