Published in final edited form as:

Patch Based Tech Med Imaging (2016). 2016 October ; 9993: 51–59. doi:10.1007/978-3-319-47118-1_7.

Multi-Atlas Based Segmentation of Brainstem Nuclei from MR Images by Deep Hyper-Graph Learning

Pei Dong¹, Yangrong Guo¹, Yue Gao², Peipeng Liang³, Yonghong Shi⁴, Qian Wang⁵, Dinggang Shen¹, and Guorong Wu¹

¹Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

²Department School of Software, Tsinghua University, Beijing, China

³Department of Radiology, Xuanwu Hospital, Capital Medical University, Beijing, China

⁴School of Basic Medical Sciences, Digital Medical Research Center, Fudan University/The Key Laboratory of MICCAI, Shanghai, China

⁵Med-X Research Institute of Shanghai Jiao Tong University, Shanghai, China

Abstract

Accurate segmentation of brainstem nuclei (red nucleus and substantia nigra) is very important in various neuroimaging applications such as deep brain stimulation and the investigation of imaging biomarkers for Parkinson's disease (PD). Due to iron deposition during aging, image contrast in the brainstem is very low in Magnetic Resonance (MR) images. Hence, the ambiguity of patch-wise similarity makes the recently successful multi-atlas patch-based label fusion methods have difficulty to perform as competitive as segmenting cortical and sub-cortical regions from MR images. To address this challenge, we propose a novel multi-atlas brainstem nuclei segmentation method using deep hyper-graph learning. Specifically, we achieve this goal in three-fold. *First*, we employ hyper-graph to combine the advantage of maintaining spatial coherence from graph-based segmentation approaches and the benefit of harnessing population priors from multi-atlas based framework. *Second*, besides using low-level image appearance, we also extract high-level context features to measure the complex patch-wise relationship. Since the context features are calculated on a tentatively estimated label probability map, we eventually turn our hyper-graph learning based label propagation into a deep and self-refining model. *Third*, since anatomical labels on some voxels (usually located in uniform regions) can be identified much more reliably than other voxels (usually located at the boundary between two regions), we allow these reliable voxels to propagate their labels to the nearby difficult-to-label voxels. Such hierarchical strategy makes our proposed label fusion method deep and dynamic. We evaluate our proposed label fusion method in segmenting substantia nigra (SN) and red nucleus (RN) from 3.0 T MR images, where our proposed method achieves significant improvement over the state-of-the-art label fusion methods.

1 Introduction

The brainstem nuclei consist of two important gray matter structures: substantia nigra (SN) and red nucleus (RN). Automatic and accurate segmentation of these two structures has high impact in clinical practice, such as deep brain stimulation and current neuroimaging studies, which aim to find robust imaging biomarkers to quantitatively measure subtle and complex structural/functional differences related to Parkinson's disease (PD). Although morphological patterns of SN and RN are more observable in functional neuro-imaging (such as SPECT) and new quantitative MRI [1], segmenting brainstem nuclei from regular 1.5 T/3.0 T T1-weighted MR images is attractive in terms of healthcare cost and the availability of imaging instruments.

Segmentation of brainstem nuclei from MR images is challenging due to (1) large shape variations across individuals and (2) low image contrast at the brainstem. To address the first challenge, multi-atlas patch-based segmentation framework [2–4] is an effective solution, which assumes two voxels are likely to have the same anatomical label if their patch-wise appearances are similar. However, conventional patch-based methods have limited power to deal with the second challenge since low image contrast makes patch-wise similarity less reliable to guide label propagation from atlas images to the target subject image. Intuitively, the first remedy is to find other complementary feature representations, which can alleviate the issue of poor image quality in brainstem area. Second, as demonstrated in graph-cut approaches [5], optimizing label fusion toward the consensus of labeling results among a group of similar voxels can reduce the risk of mis-labeling compared to the voxel-by-voxel manner used in the current multi-atlas patch-based methods. In light of this, we reckon that the integration of multi-atlas segmentation framework and graph-cut technique is a potential solution to alleviate the second issue.

To that end, we introduce hyper-graph to leverage a novel multi-atlas brainstem nuclei segmentation method through deep hyper-graph learning. Different from the conventional graph-based method that can only establish a pairwise relationship on a graph edge by connecting two vertices at a time, the hyper-graph can naturally capture group-wise relationship on a hyper-edge by linking multiple vertices [6]. Here in this paper, we use hyper-edge to partition the target subject image into a group of voxels, where subject voxels in the same hyper-edge have similar local anatomy. Meanwhile, we group each subject voxel and its matched voxels from atlas images into the same hyper-edge. Thus, the high-order information (beyond the conventional voxel-to-voxel relationship) within each hyper-edge encodes not only the multiple subject-to-atlas relationships (*i.e.*, using the power of multi-atlas segmentation framework) but also spatial coherence constraint (*i.e.*, taking the advantage of graph-based segmentation method). Given the constructed hyper-graph, the label fusion is not confined in each subject voxel separately. Instead, hyper-graph learning, dedicated to minimizing the label discrepancy within each hyper-edge, is deployed to propagate the known anatomical labels from atlas images to the target subject image domain until all subject voxels reach the unanimous agreement on label fusion.

In order to further improve the accuracy and robustness, we propose deep hyper-graph learning for nuclei segmentation by using high-level context features and hierarchical label

fusion strategy. Specifically, context features, which encode the spatial relationship of one structure to other structures, provide a complementary source of information to characterize the local anatomy in low signal-to-noise ratio regions. Since the contexts features are calculated from the tentatively estimated label probability map, we have the chance to refine the hyper-edges constructed using context features, and then iteratively correct the possible mis-labeling in hyper-graph learning. Moreover, enlightened by the hierarchical label fusion strategy in [7], we hierarchically select a set of the reliable voxels with high labeling confidence to propagate their confidently-determined labels to the nearby difficult-to-label voxels. Our proposed method has been comprehensively evaluated in segmenting SN and RN regions from T1-weighted MR images. The segmentation results show significant improvement in labeling accuracy, compared to the counterpart state-of-the-art methods [2, 4].

2 Method

The first step in our method is to register N atlas images $\{I_1, I_2, \dots, I_N\}$ to the target subject image S . After that, all voxels, including all subject image voxels $P = \{p_i | i = 1, \dots, |P|\}$ and voxels all atlas images $Q = \{q_j | j = 1, \dots, |Q|\}$, are treated as the vertices in the hyper-graph. In the following text, the subscripts i and j represent the indices for subject image voxel and atlas image voxel, respectively

2.1 Unanimous Label Fusion by Hyper-Graph Learning

Conventional multi-atlas patch-based label fusion methods find the label for each subject image voxel $p_c \in P$ by comparing only the patch-wise similarity *w.r.t.* all possible atlas image voxels. The information flow in conventional methods can be illustrated by the simple graph shown in the middle-right of Fig. 1. In this example, three atlas voxels $q_1 \sim q_3$ (red circles) are matched with p_c (red box). Note that there is no connection among subject image voxels, although such neighboring information in the subject image is useful to maintain the spatial coherence.

Hyper-Edge Construction—In light of this, we propose to use hyper-graph to reveal such complex relationship and achieve unanimous label propagation for all subject voxels. Compared to the conventional graph edge which can only connect two graph vertices, each hyper-edge e allows to accommodate more than two vertices. In order to address the difficulty of low image contrast, we take advantage of hyper-edge to find more heuristics for label fusion in two folds: **(1) High-order self-similarity relationship**. Intuitively, we allow each subject image voxel p_c to find a set of spatially-close and anatomically-similar voxels in the subject image domain. As the purple dot contour shown in the bottom-right of Fig. 1, four subject voxels $p_1 \sim p_4$ fall in the same hyper-edge centered at p_c . Thus, we can partition the whole subject image into a group of voxels which is essentially one hyper-edge describing the high-order self-similarity relationship (beyond two voxels). **(2) Multi-channel subject-to-atlas relationship**. Existing methods can only use pre-defined similarity measurement, which is not adaptive to measure the complicated relationship between subject and atlas voxels at each location. In contrast, hyper-graph allows us to efficiently integrate diverse information from multiple channels by using various hyper-edges. For example, we

can find a group of atlas image voxels (q_1 and q_2) which have similar patch-wise appearance as p_c and then form a hyper-edge as displayed in black dot contour in the bottom-right of Fig. 1. Given the tentative label probability map, we further extract context features for each voxel and form another hyper-edge in terms of context features centered at p_c , as shown in the bottom-right of Fig. 1 where p_c , q_2 , and q_3 are included in the blue dot contour. Thus, many hyper-edges can be built to describe the complete characteristics of anatomical structure from different views. Compared to the simple information flow used in conventional label fusion methods, the hyper-graph (bottom-right of Fig. 1) contains much richer information to drive the label fusion.

Hyper-Graph Construction—In our method, the hyper-graph is denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \mathcal{V}_P \cup \mathcal{V}_Q$ is the vertices set which includes all subject voxels P and atlas voxels Q . $\mathcal{E}$ denotes the set of hyper-edges constructed above. The information conveyed by various hyper-edges is encoded by an incidence matrix $\mathbf{H}_{|\mathcal{V}| \times |\mathcal{E}|}$. Each entry $h(v, e)$ in $\mathbf{H}$ measures the affinity degree between each vertex $v \in \mathcal{V}$ and the underlying center p_c in the hyper-edge $e \in \mathcal{E}$. Take the image appearance features as example, the degree of $h(v, e)$ can be defined as:

$$h(v, e) = \begin{cases} \exp\left(-\frac{\Omega(\kappa(v)) - \Omega(p_c)^2}{\sigma^2}\right) & \text{if } v \in e \\ 0 & \text{if } v \notin e \end{cases} \quad (1)$$

where Ω is the operator of extracting intensity image patch at a given center. κ is an index function that maps each vertex $v \in \mathcal{V}$ to the spatial location in particular image. σ is used to control the strength of decay. Furthermore, we construct a $|\mathcal{E}| \times |\mathcal{E}|$ diagonal matrix $\mathbf{W}$ where each diagonal element $w(e)$ indicates the weight for particular hyper-edge e . Since there is no prior information on hyper-edges, we set $w(e) = 1$ ($\forall e \in \mathcal{E}$). Hence $\mathbf{W}$ is eventually an identity matrix. The degree of each vertex v is defined as $d(v) = \sum_{e \in \mathcal{E}} w(e)h(v, e)$, and the degree of hyper-edge is defined as $\delta(e) = \sum_{v \in \mathcal{V}} h(v, e)$. Thus, two diagonal matrices $\mathbf{D}_v$ and $\mathbf{D}_e$ can be formed with each entry along the diagonal using the vertex degree and hyper-edge degree, respectively. Hyper-Graph Learning. After establishing the relationships among all the voxels through hyper-graph, the next step is to propagate anatomical labels from atlas vertices $\mathcal{V}_Q$ to subject image vertices $\mathcal{V}_P$. For convenience, let y_j denote the label on the atlas vertex $v_j \in \mathcal{V}_Q$, which has two values, ‘-1’ for background and ‘1’ of foreground (the underlying structure). We initially set the label y_i on the subject image vertex $v_i \in \mathcal{V}_P$ to ‘0’, since the anatomical labels on the subject image are unknown. Furthermore, we arrange labels on all vertices into a column vector $\vec{Y} = \left[[y_i]_{i=1}^{|P|}, [y_j]_{j=1}^{|Q|} \right]^T$.

The goal of label fusion is to jointly optimize relevance values $\vec{F} = \left[[f_i]_{i=1}^{|P|}, [f_j]_{j=1}^{|Q|} \right]^T$ for all vertices, regardless from atlas or subject images, which suggests the preference of

choosing background (<0) or foreground (>0). Thus, our label fusion method falls into the semi-supervised hyper-graph learning framework, i.e., the to-be-determined labels on the vertices from the subject image are influenced by the connected vertices from atlas images with known labels as well as subject voxels under labeling. The principle to propagate the labels is that the vertexes sitting in the same hyper-edge are likely to have the same label, by also keeping the minimal discrepancy of labels on the vertexes with known labels before and after label propagation. The objective function of hyper-graph learning is defined as:

$$\arg \min_{\vec{F}} \left\{ \|\vec{Y} - \vec{F}\|_2^2 + \lambda \cdot \Phi(\vec{F}, \mathbf{H}, \mathbf{W}, \mathbf{D}_e, \mathbf{D}_v) \right\} \quad (2)$$

The first term in Eq. 2 is called the fitting term, which requires the estimated relevance vector $\vec{F}$ and the initialization vector $\vec{Y}$ should have minimal discrepancy before and after label propagation.

The second term $\Phi(\cdot)$ is the graph balance term [8], defined as:

$$\Phi(\vec{F}, \mathbf{H}, \mathbf{W}, \mathbf{D}_e, \mathbf{D}_v) = \frac{1}{2} \sum_{e \in \mathcal{E}} \sum_{v, v' \subseteq e} \frac{w(e)h(v, e)h(v', e)}{\delta(e)} \left(\frac{f(v)}{\sqrt{d(v)}} - \frac{f(v')}{\sqrt{d(v')}} \right)^2 \quad (3)$$

which holds the constraint in each hyper-edge that two vertices v and v' in the same hyper-edge should have similar relevance value. The λ is a positive weighing parameter between the two terms. We can rewrite the graph balance term into

$\Phi(\vec{F}, \mathbf{H}, \mathbf{W}, \mathbf{D}_e, \mathbf{D}_v) = \vec{F}^T (\mathbf{I} - \mathbf{\Theta}) \vec{F} = \vec{F}^T \mathbf{\Delta} \vec{F}$, Where $\mathbf{\Theta} = \mathbf{D}_v^{-\frac{1}{2}} \mathbf{H} \mathbf{W} \mathbf{D}_e^{-1} \mathbf{H}^T \mathbf{D}_v^{-\frac{1}{2}}$, $\mathbf{I}$ is the identity matrix, and $\mathbf{\Delta}$ acts as a normalized hyper-graph Laplacian matrix. By differentiating the objective function with respect to $\vec{F}$, the optimal $\hat{\vec{F}}$ can be computed by:

$$\hat{\vec{F}} = (\mathbf{I} + \lambda(\mathbf{I} - \mathbf{\Theta}))^{-1} \vec{Y} \quad (4)$$

Given the estimated $\hat{\vec{F}}$, the anatomical label on the subject image can be easily determined by the sign of the relevance value.

2.2 Hierarchical Label Propagation Mechanism

Some subject image voxels might have the higher confidence to be labeled than others due to better alignment of atlas images or locating in the uniform regions in the image. In light of this, we partition the vertices in $\mathcal{V}_p$ into two groups based on their difficulties in labeling. To achieve it, we use an existing label fusion method, such as the majority voting, to predict the label on each subject voxel as well as the confidence value in terms of the voting predominance. If the influence for voting one label dominates other labels, we assume the estimated label on the underlying subject voxel can be treated as ground-truth. Otherwise,

we need to confirm the underlying voxel label by other heuristics. The advantage is that it allows label propagation from *not only* the atlas images *but also* some reliable regions of subject image, which are more specific to label fusion of subject image. After label fusion by hyper-graph learning, the refined labeling result brings more subject voxels with high labeling confidence. By iteratively selecting more and more subject vertices with high confidence and performing label fusion, we can gradually improve the label fusion result.

2.3 Deep Hyper-Graph Learning by Using High-Level Context Features

In Sect. 2.1, we only use the image appearance information to construct the hyper-graph. Due to low image contrast, low-level features derived from image intensity are not sufficient to steer hyper-graph learning. Other high-level features are of great necessity to alleviate the issue of poor image quality. To this end, we resort to context features which can encode spatial relationship of one structure to other structures. The anatomical structures are labeled with different colors, as shown in the blue box of Fig. 2. Thus, for each voxel, the label information from other surrounding voxels provides an additional clue to identify the underlying voxel in terms of its relationship to other structures. In order to make the context features more robust and efficient, we further construct the context feature profile (black dots in the red boxes of Fig. 2) in a multi-resolution manner, where we densely sample nearby the center (to represent the characteristics of local anatomy) and gradually reduce sampling outward (to capture the global scale information).

Given the context features, we can refine the segmentation result with more accurate self-similarity and atlas-subject relationship revealed in the hyper-graph. Since the context features of the subject image are extracted from the tentatively refined label probability map, we can turn our label fusion method into a deep model. As shown in Fig. 2, the refined context features lead to more accurate segmentation results, as the label probability map (displayed in darker blue boxes) becomes clearer and clearer. In summary, by repeating (1) construction of hyper-graph with both low-level image features and high-level context features, (2) label fusion via hyper-graph learning and (3) update of the context features, we have the chance to correct mislabeling and hence improve the labeling accuracy.

3 Experiments

In the experiments, 11 MR images of PD subjects are used to evaluate the proposed label fusion method. Each subject has the SN and RN structures manually labeled by two radiologists. The image voxel size is $1 \times 1 \times 1 \text{ mm}^3$. For all the experiments, the patch size is set to $5 \times 5 \times 5 \text{ mm}^3$. The search neighborhood in constructing hyper-edges is set to $3 \times 3 \times 3 \text{ mm}^3$. Parameter λ in Eq. 2 is 0.1.

To evaluate the performance of the proposed method, we adopted the leave-one-out cross-validation, where each subject was in turn used as a subject image and the remaining 10 subjects were used as atlas images. Our proposed deep hyper-graph patch labeling (DHPL) method is compared with two state-of-the-art methods: nonlocal mean patch-based label fusion method (NLM) [2], and sparse patch-based label fusion method (SPBL) [8]. Note, we compare our method with the best segmentation results by these counterpart methods.

Table 1 shows the mean and standard deviation of Dice ratios (DR) between manual segmentation and the estimated segmentation by three methods. Compared to the SPBL, our method can obtain overall 2.2 % and 1.3 % improvement ratio for SN and RN, respectively. Paired t-test further confirmed that our full method (DHPL) achieves the significant improvement in terms of DR over the two counterpart methods. Figure 3 further shows the average symmetric surface distance (ASSD) between the ground truth and estimated segmentation result by the three methods. Based on the color map shown in the right of Fig. 3, the segmentation result by our DHPL method is much closer to the ground truth than any other methods.

4 Conclusion

In this paper, we proposed a novel multi-atlas label fusion framework for brainstem nuclei segmentation for MR images using deep hyper-graph learning. The novelty of our method lies in three ways. First, we employed hyper-graph learning method to label anatomical regions, by taking its advantages of both maintaining the spatial consistency and coupling population priors from multiple atlases. Second, we harnessed high-level context features to steer the hyper-graph learning in order to overcome the poor image appearance. Here, we implemented our method in a deep and self-refining model by extracting context features from the tentatively estimated label probability map. Third, we allowed more and more reliable predicted subject voxel labels to guide label propagation to those difficult-to-label voxels, making our method *not only* deep *but also* dynamic. Experimental results show significantly enhanced better segmentation results using our proposed method compared to the state-of-the-art patch-based label fusion methods.

References

1. Ziegler DA, Augustinack JC. Harnessing advances in structural MRI to enhance research on Parkinson's disease. *Imaging Med.* 2013; 5:91–94. [PubMed: 23687517]
2. Coupé P, Manjón JV, Fonov V, Pruessner J, Robles M, Collins DL. Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. *NeuroImage.* 2011; 54:940–954. [PubMed: 20851199]
3. Wu G, Kim M, Sanroma G, Wang Q, Munsell B, Shen D. Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition. *NeuroImage.* 2015; 106:34–46. [PubMed: 25463474]
4. Zhang, D., Guo, Q., Wu, G., Shen, D. Sparse patch-based label fusion for multi-atlas segmentation. In: Yap, P.-T., Liu, T., Shen, D., Westin, C.-F., Shen, L., editors. *MBIA 2012 LNCS.* Vol. 7509. Springer; Heidelberg: 2012. p. 94–102.
5. Shi J, Malik J. Normalized cuts and image segmentation. *IEEE Trans Pattern Anal Mach Intell.* 2000; 22:888–905.
6. Dong, P., Guo, Y., Shen, D., Wu, G. Multi-atlas and multi-modal hippocampus segmentation for infant MR brain images by propagating anatomical labels on hypergraph. In: Wu, G., Coupé, P., Zhan, Y., Munsell, B., Rueckert, D., editors. *Patch-MI 2015 LNCS.* Vol. 9467. Springer; Heidelberg: 2015. p. 188–196.
7. Zhang, D., Wu, G., Jia, H., Shen, D. Confidence-guided sequential label fusion for multi-atlas based segmentation. In: Fichtinger, G., Martel, A., Peters, T., editors. *MICCAI 2011, Part III LNCS.* Vol. 6893. Springer; Heidelberg: 2011. p. 643–650.
8. Zhou D, Huang J, Schölkopf B. Learning with hypergraphs: clustering, classification, and embedding. *Advances in NIPS.* 2006; 19:1601–1608.

Fig. 1.

Comparison of simple graph model in conventional methods and hyper-graph used in our method (with much richer information).

Fig. 2.

Illustration of the deep hyper-graph learning leveraged by context features.

Fig. 3.

Visual comparison of segmentation results by three label fusion methods on a typical subject. Automatic segmentations are shown in red contours and manual segmentations are shown in yellow contours in the bottom row. (Color figure online)

Table 1

The mean and standard deviation of the Dice ratio by three methods.

	NLM	SPBL	DHPL
SN	72.2 ± 2.7 % *	73.0 ± 1.9 % *	75.2 ± 2.2 %
RN	82.2 ± 2.8 % *	83.0 ± 1.8 % *	84.3 ± 1.9 %

* Indicates significant improvement of our DHPL over other counterpart methods (p -value < 0.05)