HIGH-ORDER BOLTZMANN MACHINE-BASED UNSUPERVISED FEATURE LEARNING FOR MULTI-ATLAS SEGMENTATION

Liang Sun, Wei Shao, Daoqiang Zhang

Dept. of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016,China

ABSTRACT

Multi-atlas based label fusion methods have been successfully used for medical image segmentation. In the field of brain region segmentation, multi-atlas based methods propagate labels from multiple atlases to target image by the similarity between patches in target image and atlases. Most of existing multi-atlas based methods usually use intensity feature, which is hard to capture high-order information in brain images. In light of this, in this paper, we endeavor to apply high-order restricted Boltzmann machines to represent brain images and use the learnt feature for brain region of interesting (ROIs) segmentation. Specifically, we firstly capture the covariance and the mean information from patches by high-order Boltzmann Machine. Then, we propagate the label by the similarity of the learnt high-order features. We validate our feature learning method on two well-known label fusion methods e.g., local-weighted voting (LWV) and non-local mean patch-based method (PBM). Experimental results on the NIREP dataset demonstrate that our method can improve the performance of both LWV and PBM by using the high-order features.

Index Terms— Multi-atlas, Label fusion, High-order restricted Boltzmann machines

1. INTRODUCTION

Automatic and accurate segmentation of Magnetic resonance (MR) images is a critical step for pathology detection and brain parcellation, etc. For example, for the task of diagnosing Alzheimers disease, it needs medical expert to segment hippocampus, which is highly related to this disease [1, 2], from the MR images at the first step. However, it is time-consuming with high labor intensity. Accordingly, it is necessary to develop an accurate and automatic method for brain segmentation.

Recently, multi-atlas based segmentation methods have gained great success in segmenting brain into difference ROIs [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. Generally, there are two main steps (i.e., image registration and label fusion) for multi-atlas segmentation. Specifically, MR images will be warped onto a common space in the first step, and then labels from different atlases will be propagated to the target image to obtain its corresponding final labels in the next step.

During last decade, numerous label fusion strategies have been proposed for multi-atlas based segmentation. Specifically, majority voting (MV) [3] is the simplest one, by which each atlas image is viewed equally when assigning its label to the target image. More advanced strategy like local-weighted voting (LWV) [4] method considers the patch-wise similarity between the target and each atlas as the voting weight. Moreover, in order to alleviate possible registration errors, the non-local mean patch-based method (PBM) [9, 10] focuses on propagating labels by using the similarity between the patch from target image and atlas image within a certain neighborhood. The results indicate that PBM can improve the accuracy and robustness of the segmentation performance. In addition, the sparse learning method (i.e., SPBM) [11, 12] is also used for label fusion, where a small number of patches from multiple atlases are used for reconstructing the target patch.

Although much progress has been achieved, the above methods only use the intensity information to describe MR images, which neglect to take high-order information into consideration. The mean-covariances Restricted Boltzmann Machine model (mcRBM) [13], which can effectively capture high-order information, such as mean and covariance information from images, simultaneously. Specifically, the mcRBM has two sets of hidden units, one represents the patch intensities (i.e., mean information) and the other one represents pair-wise dependencies that depict the high order information between patches intensities (i.e., covariance information). In our work, we firstly apply mcRBM model to learn mean and covariance information from the image patches, and then use the learnt high-order feature for multi-atlas label fusion. We validate the learnt feature on two well-known label fusion methods including LWV and PBM. Experimental results on the NIREP data sets show our method can improve the performance of both LWV and PBM by the learnt feature.

The remainder of this paper is organized as follows. In section 2, we present the feature learning process of mcRBM and the label fusion method. In Section 3, we compare our

This study was supported by NSFC (Nos. 61422204, 61473149) and the NUAA Fundamental Research Funds (No. NE2013105).

proposed method with LWV and PBM . Finally, a brief conclusion is given in section 4.

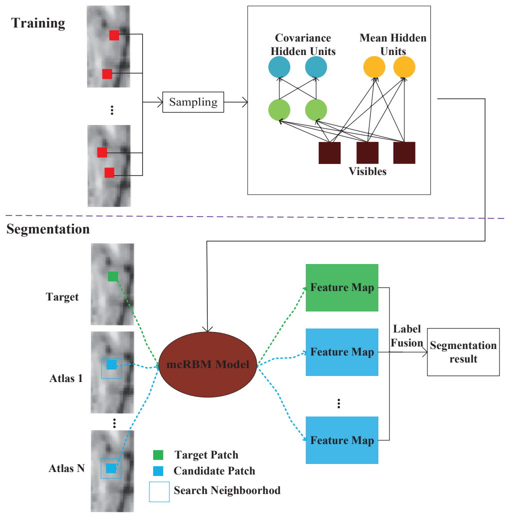

Fig. 1. Overview of our method.1) In the training stage, the red squares represent the patches random sample from MRI to train mcRBM model. 2) In the multi-atlas segmentation stage, the green squares represent the target patch and the blue patches represent the candidate patches selected from the atlases. And the blue rectangle represent the search neighborhood.

2. METHOD

We show the flowchart of our method in Fig.1. Specifically, it can be divided into two stages (e.g., training stage and segmentation stage). In the training stage, we random sample patches from atlases and target image to train the mcRBM model offline. In the segmentation stage, we firstly convert the intensity feature to the new feature space by the derived mcRBM model in the training stage, and then adopt the label fusion methods (e.g., PBM, LWV) to get the final label map of the brain image. We will present the feature learning process of mcRBM and the label fusion method in Section 2.1 and Section 2.2, respectively.

2.1. High-order feature learning

The mcRBM model [13] has two sets of hidden units, one represents the mean information of patch and the other represents the covariance information of patch. Thus, the hidden units can be used to capture both the mean and the pair-wise dependencies between voxels. The mean intensities of the visible units energy function is defined as:

$$E^m(v, h^m) = -\sum_{j=1}^{M} \sum_{i=1}^{D} W_{ij}\, h_j^m\, v_i - \sum_{j=1}^{M} b_j^m\, h_j^m \tag{1}$$

where h_j^m represent the *j*-th mean hidden unit, v_i is the *i*-th visible unit, b_j^m is biases and $W_{i,j}$is the connection weight between h_j^m and v_i. The covariance of the visible unites energy function is defined as:

$$\begin{aligned} E^c(v, h^c) = -\tfrac{1}{2} \textstyle\sum_{f=1}^{F} \sum_{k=1}^{N} P_{fk}\, h_k^c \left(\sum_{i=1}^{D} C_{if} v_i\right)^2 \\ - \textstyle\sum_{k=1}^{N} b_k^c\, h_k^c \end{aligned} \tag{2}$$

where h_k^c represents the *k*-th covariance hidden unit, C is the connection weight matrix between visible units and filters, and P is the connection weight matrix between filters and hidden units. The final energy function is:

$$E^m(v, h^c, h^m) = E^m(v, h^m) + E^c(v, h^c) \tag{3}$$

The probability density function can be defined as:

$$P(v, h^c, h^m) = \frac{1}{Z} \exp^{E^m(v, h^c, h^m)} \tag{4}$$

with

$$Z = \sum_{h^c, h^m} E^m(v, h^c, h^m)$$

where Z is a normalized factor. The marginal distribution over the visible units is $P(v) = \sum_{h^c, h^m} P(v, h^c, h^m)$. We can solve this optimization problem by applying Contrastive Divergence method[14].

2.2. Patch-based method for label fusion

Here, we use T to denote the label of the target image. Label fusion aims to determine the label map L_T for the target image. We firstly register each atlas image and label maps onto the target image space. We use $A = \{A_s\, |s = 1, ..., N\}$ and $L = \{L_s\, |s = 1, ..., N\}$ to denote the N atlases and label maps ,and use $P_T(y)$ and $P_{A_S}(x)$ to denote the patch centered at the voxel y in the target image T and the patch centered at voxel x in the atlas A_s, respectively. In addition, we denote the neighborhood of voxel y on the atlases A as $N_A(y)$.

In the non-local strategy, the patch-based methods seek multiple candidates in a certain neighborhood centered at the target voxels. The multi-atlas patch-based method assume that the target voxels should have the same label as the atlas voxels if the local tissue shape or appearance is similar, In PBM [9], the voting weights are calculated as follows:

$$w(y, x_{s,j}) = \exp \frac{-||I(y) - I(x_{s,j})||_2^2}{\delta(y)} \tag{5}$$

with

$$\delta(y) = \underset{x_{s,j} \in N_A(y)}{\arg\min} \; ||I(y) - I(x_{s,j})||_2 + \varepsilon$$

where $I(y)$ and $I(x_{s,j})$ represent the normalized intensity of the patches $P_T(y)$ and $P_{A_S}(x_{s,j})$. $||\cdot||_2$ is the L2 norm. ε is a small constant. In our method we replace intensity with the high-order descriptor. For all voxels y of the target image to be segmented, the label of voxel is determined by a weighted label fusion of all labeled voxels inside the neighborhood of voxel y:

$$l(y) = \frac{\sum_{s=1}^{N} \sum_{j \in N_A(y)} w(y, x_{s,j}) \, l_{s,j}}{\sum_{s=1}^{N} \sum_{j \in N_A(y)} w(y, x_{s,j})} \tag{6}$$

where $l_{s,j}$ is the label given by the expert to voxel $x_{s,j}$at the *j*th voxel in atlas A_s. $w(y, x_{s,j})$ is a weight between voxel y and $x_{s,j}$ calculate by Eq.5. The final label is assigned to 1 or 0 by

$$L(y) = \begin{cases} 1 & l(y) \geq 0.5 \\ 0 & l(y) < 0.5 \end{cases} \tag{7}$$

3. EXPERIMENTS

In this section, we evaluate the performance of the proposed label fusion method on segmenting brain anatomical structure, i.e., regions of interest (ROIs), from the MR brain image of NIREP database. We validate the learnt feature on two well-known label fusion methods including LWV and PBM.

3.1. Dataset and experimental settings

The NIREP dataset [15] consist of 16 T1-weighted MR images, including 8 normal adult males and 8 females. The 16 MR images have been manually segmented into 32 ROIs. The MR images were obtained in a General Electric Signa scanner operating at 1.5 Tesla, using the following protocol: SPGR/50, TR24, TE 7, NEX 1 matrix 256 192, FOV 24 cm. For each of the ROIs, a Leave-One-Out cross-validation is performed to test the segmentation on each LOO fold, 15 of 16 subjects are used as atlases and aligned onto the remaining image (used as the target image). We use the Dice ratio to assess label accuracy which measures the degree of overlap, defined as:

$$Dice(A, B) = \frac{2 \, |A \cap B|}{|A| + |B|} \tag{8}$$

where the $\cap$ denotes the overlap between automatic segmentations and ground truth, and $|\cdot|$ denotes the number of voxels of the ROI.

We use the terms of mc-LWV and mc-PBM to denote the methods of LWV and PBM, by which we apply mcRBM to obtain the high-order feature, respectively. We perform a preselection of the patches to reduce the computational time according to the similarity of the two patches. There is a common parameter in our experiment, i.e., the size of 7×7×7 neighborhood search region is used for our experiments. And in our experiments we fix patch size as 5×5×5 voxels for all methods. We train mcRBM model with 256 factors, 256 covariance hidden units and 100 mean hidden units.

3.2. Experimental results on the NIREP dataset

We compare the segmentation results of different multi-atlas label fusion algorithms on NIREP dataset. There are 32 ROIs in the NIREP dataset, for simplicity, we treat them independently. Thus we have 32 independent binary segmentation. For each segmentation, we set the label of a voxel as 1 if it belongs to the ROI, and 0 otherwise. Table 1 gives the average dice of the 32 ROIs. The mc-LWV and mc-PBM achieve better segmentation results than LWV and PBM, respectively. Compared with LWV, mc-LWV improve the average dice ratio(mean±standard deviation) from (61.58±7.22)% to (62.70±7.42)%. Compared with PBM, mc-PBM improves the average dice ratio from (75.19±5.28)% to (76.73±5.39)%. We can see that the feature learning by mcRBM model can improve the segmentation results. In Fig.2 we visually plot the segmentation results of the LWV, mc-LWV, PBM and mc-PBM on *L insula gyrus*. We also show the ground truth and original brain image for comparison. As can see from Fig.2, mc-LWV, mc-PBM achieve a better visual quality of segmentation results than LWV and PBM, respectively. In supplemental material, we shows the segmentation results on the 32 ROIs using LWV, mc-LWV, PBM and mc-PBM. our proposed feature learning methods can be improve the performance for both LWV and PBM on all ROIs, respectively.

Table 1. Segmentation results of LWV, mc-LWV, PBM and mc-PBM on NIREP dataset

Method	Mean±Standard deviation(%)
LWV	61.58±7.22
mc-LWV	**62.70±7.42**
PBM	75.19±5.28
mc-PBM	**76.73±5.39**

4. CONCLUSION

In this paper, we use mcRBM to learn high-order feature for multi-atlas segmentation. The learnt feature can both capture the information of the voxel intensities and pair-wise dependencies between voxel intensities. The segmentation results show that our proposed mc-LWV and mc-PBM can achieve up to 1.12% and 1.54% improvement over LWV and PBM on NIERP dataset, respectively. These results demonstrate the advantage of the newly learnt feature in the field of multi-atlas segmentation. In future, we plan to use this model for more segmentation task on medical image dataset and extend it to build a Deep Belief Network (DBN) for learning deep features.

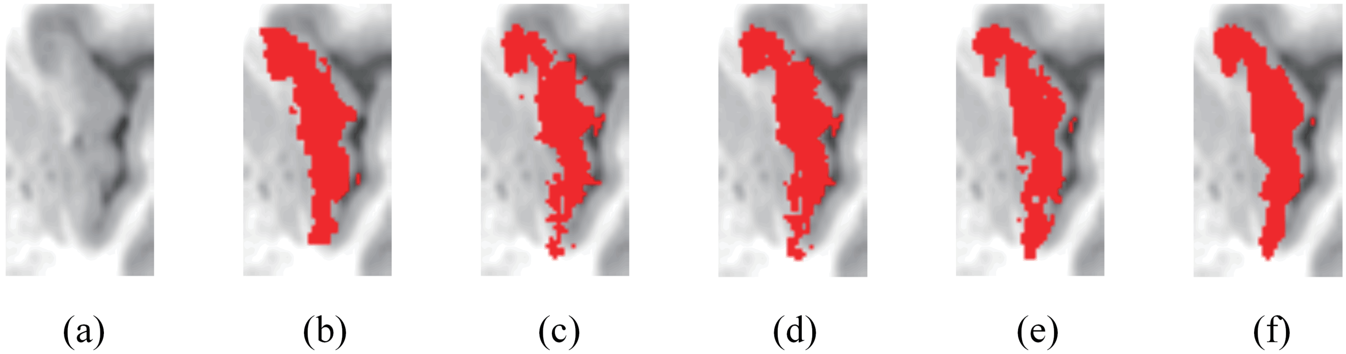

Fig. 2. Visual views on original image (a), ground truth (b) and segmentation result of LWV (c), mc-LWV(d), PBM (e) and mc-PBM (f) algorithms on *L insula gyrus*

5. REFERENCES

[1] Manhua Liu, Daoqiang Zhang, and Dinggang Shen, “Ensemble sparse classification of alzheimers disease,” *NeuroImage*, vol. 60, no. 2, pp. 1106, 2012.

[2] Chen Zu, Biao Jie, Mingxia Liu, Songcan Chen, Dinggang Shen, and Daoqiang Zhang, “Label-aligned multitask feature learning for multimodal classification of alzheimers disease and mild cognitive impairment,” *Brain Imaging and Behavior*, 2015.

[3] Rolf A Heckemann, Joseph V Hajnal, Paul Aljabar, Daniel Rueckert, and Alexander Hammers, “Automatic anatomical brain mri segmentation combining label propagation and decision fusion.,” *NeuroImage*, vol. 33, no. 1, pp. 115–126, 2006.

[4] Xabier Artaechevarria, Arrate Munozbarrutia, and Carlos Ortizdesolorzano, “Combination strategies in multiatlas image segmentation: Application to brain mr data,” *IEEE Transactions on Medical Imaging*, vol. 28, no. 8, pp. 1266, 2009.

[5] Suyash P. Awate and Ross T. Whitaker, “Multiatlas segmentation as nonparametric regression,” *IEEE Transactions on Medical Imaging*, vol. 33, no. 9, pp. 1803, 2014.

[6] Guorong Wu, Minjeong Kim, Gerard Sanroma, Qian Wang, Brent C Munsell, and Dinggang Shen, “Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition,” *NeuroImage*, vol. 106, pp. 34, 2015.

[7] Thomas R Langerak, U A Van Der Heide, Alexis N T J Kotte, Max A Viergever, M Van Vulpen, and J P W Pluim, “Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (simple),” *IEEE Transactions on Medical Imaging*, vol. 29, no. 12, pp. 2000–2008, 2010.

[8] Haidong Wang, Jung W Suh, S R Das, John Pluta, Caryne Craige, and Paul A Yushkevich, “Multi-atlas segmentation with joint label fusion,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 3, pp. 611–623, 2013.

[9] Pierrick Coupe, Jose V Manjon, Vladimir Fonov, Jens C Pruessner, Montserrat Robles, and D Louis Collins, “Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation.,” *NeuroImage*, vol. 54, no. 2, pp. 940–954, 2011.

[10] F Rousseau, Piotr A Habas, and Colin Studholme, “A supervised patch-based approach for human brain labeling,” *IEEE Transactions on Medical Imaging*, vol. 30, no. 10, pp. 1852–1862, 2011.

[11] Tong Tong, Robin Wolz, Pierrick Coupe, Joseph V Hajnal, and Daniel Rueckert, “Segmentation of mr images via discriminative dictionary learning and sparse coding: Application to hippocampus labeling,” *NeuroImage*, vol. 76, no. 1, 2013.

[12] Daoqiang Zhang, Qimiao Guo, Guorong Wu, and Dinggang Shen, “Sparse patch-based label fusion for multiatlas segmentation,” in *International Workshop on Multimodal Brain Image Analysis*. Springer, 2012, pp. 94–102.

[13] Geoffrey E Hinton et al., “Modeling pixel means and covariances using factorized third-order boltzmann machines,” in *CVPR2010*. IEEE, 2010, pp. 2551–2558.

[14] Geoffrey E Hinton, “Training products of experts by minimizing contrastive divergence,” *Neural Computation*, vol. 14, no. 8, pp. 1771–1800, 2002.

[15] Gary E Christensen, Geng, et al., “Introduction to the non-rigid image registration evaluation project (nirep),” in *International Workshop on Biomedical Image Registration*. Springer, 2006, pp. 128–135.