Published in final edited form as:

Proc IEEE Int Symp Biomed Imaging. 2015 April ; 2015: 1560–1563. doi:10.1109/ISBI.2015.7164176.

Joint Labeling Of Multiple Regions of Interest (RoIs) By Enhanced Auto Context Models

Minjeong Kim, Guorong Wu, Yanrong Guo, and Dinggang Shen

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

Abstract

Accurate segmentation of a set of regions of interest (ROIs) in the brain images is a key step in many neuroscience studies. Due to the complexity of image patterns, many learning-based segmentation methods have been proposed, including auto context model (ACM) that can capture high-level contextual information for guiding segmentation. However, since current ACM can only handle one ROI at a time, neighboring ROIs have to be labeled separately with different ACMs that are trained independently without communicating each other. To address this, we enhance the current single-ROI learning ACM to multi-ROI learning ACM for joint labeling of multiple neighboring ROIs (called *e*ACM). First, we extend current independently-trained single-ROI ACMs to a set of jointly-trained cross-ROI ACMs, by simultaneous training of ACMs for all spatially-connected ROIs to let them to share their respective intermediate outputs for coordinated labeling of each image point. Then, the context features in each ACM can capture the cross-ROI dependence information from the outputs of other ACMs that are designed for neighboring ROIs. Second, we upgrade the output labeling map of each ACM with the multi-scale representation, thus both local and global context information can be effectively used to increase the robustness in characterizing geometric relationship among neighboring ROIs. Third, we integrate ACM into a multi-atlases segmentation paradigm, for encompassing high variations among subjects. Experiments on LONI LPBA40 dataset show much better performance by our *e*ACM, compared to the conventional ACM.

Index Terms

Labeling; Auto context model (ACM)

1. Introduction

Accurate quantitative measurement of morphological structures is very important in neuroimaging-based studies. Since spatially connected or adjacent structures in the brain are highly correlated to each other in term of their brain functions or changes due to diseases, it is demanding to segment them jointly from brain images. However, it is often difficult to accurately label voxels in the boundaries of neighboring structures, due to the complex image patterns.

Among many automatic ROI labeling methods [1-7], auto-context model (ACM) [7] is one of state-of-the-art learning-based methods, which can use not only image appearance but context features for guiding labeling. In the beginning of training stage, image appearance

features are used to train a classifier and estimate a probability map, for evaluating the probability of each image point belonging to the target ROI. Then, the context information can be derived from this tentative probability map to train next classifier by providing the additional geometric context heuristics. By alternating 1) the training of next classifier from both image appearance and the latest context features, 2) the estimation of probability map by the current learned classifier for updating context features, a sequence of classifiers can be learned to segment the target ROI from the training images. In the testing stage, given a new subject, the classification probability map will be iteratively refined via a sequence of those trained classifiers, to obtain the final segmentation.

However, similar to many other methods, current ACM is limited for segmenting one ROI at a time. In this paper, we extend the conventional single-ROI learning ACM to the multi-ROI joint learning ACM. *First*, we simultaneously train ACMs and coordinate all ACMs in predicting the label on each point, where each ACM is trained for a particular ROI. For this, we allow the ACMs to share their respective intermediate outputs during the training. Thus, the context features in each ACM can capture not only the within-ROI relation information from its own outputs, but the cross-ROI dependence information from the outputs of other ACMs that are designed for the neighboring ROIs. *Second*, we further upgrade the output labeling map of each ACM with the multi-scale representation, thus both local and global context information can be effectively used to increase the robustness in characterizing geometric relationship among the neighboring ROIs. *Third*, we integrate ACM into a multi-atlases segmentation paradigm, for encompassing the high variations among different subjects.

By comparison with conventional ACM using the LONI LPBA40 dataset with 56 manual ROIs for each brain image, our *e*ACM has achieved more accurate segmentation results in terms of ROI overlap degree. It is worth noting that it is straightforward to apply our *e*ACM method to simultaneous segmentation of *other organs* in many other applications.

2. Method

The goal of ACM is to learn a set of classifiers to gradually label each point in a new target image T into either positive (e.g., object) or negative (e.g., non-object). In the following, we first briefly explain the conventional ACM in Section 2.1. Then, we present our augmented ACM with the capability of jointly handling N non-overlapping ROIs $O = \{O_n \mid n = 1, \dots, N\}$ in Section 2.2. We further enhance the robustness of our method by using multi-scale representation and multi-atlases paradigm in Section 2.3.

2.1. Conventional Single-ROI ACM

Learn a Sequence of Classifiers in the Training Stage—Given M aligned training images $I = \{I_i \mid i = 1, \dots, M\}$ and their corresponding manual label maps $L = \{L_i \mid i = 1, \dots, M\}$, we train a sequence of classifiers from local image appearance and context features to predict whether a particular point x belongs to a certain ROI O_n ($n = 1, \dots, N$). Specifically, for each point x , its local image appearance features can be extracted in a certain neighborhood, which include (1) image intensities and statistics (e.g., intensity mean and variance) within the local neighborhood, (2) Haar features at different scales, and (3) texture

features (e.g., image contrast and entropy). Besides using these image appearance features, context features importantly used in ACM. Since each probability map $P_{i,n}(x)$ (pink boxes in Fig. 1(a)), predicted by the classifiers of each ROI, indicates the likelihood of each point x in the image I_i belonging to the ROI O_n , the profile of such probabilities centered at point x can provide shape prior information. Before computing the context features, we first define the sampling pattern $\delta = \{\delta_q \mid q = 1, \dots, Q\}$ upon each classification probability map, consisting of a sequence of displacement vectors δ_q w.r.t. the center point x with the number of sampling positions Q . A typical sampling pattern is shown in Fig. 1(a), where the yellow box denotes the sampling location at the point x and the blue boxes denote the sampling locations. Thus, the ROI-specific context features, $\theta_{i,n}(x)$ (w.r.t. the n -th ROI O_n) on point x of the training image I_i , can be eventually collected as a stack of classification probabilities by sequentially visiting each location in the sampling pattern, i.e., $\theta_{i,n}(x) = [P_{i,n}(x + \delta_1), \dots, P_{i,n}(x + \delta_q), \dots, P_{i,n}(x + \delta_Q)]$.

Since ACM trains the classifier iteratively, we use k ($k = 1, \dots, K$) to denote the k -th iteration. In the k -th iteration, image appearance and ROI-specific context features $\theta_{i,n}(x)$ on each point x of training image are extracted from the intensity images I_i and the classification probability map $P_{i,n}^{(k-1)}$ in the $(k-1)$ -th iteration, respectively. After repeating the same feature extraction procedure across all training images, we can train classifiers to learn the nonlinear mapping from the appearance and context features to the manual label maps L .

Then, we can generate a new classification probability map $P_{i,n}^{(k)}$ of ROI O_n for the next iteration, until convergence. It is apparent the conventional ACM has limitation for handling multiple ROIs, since (1) the context feature $\theta_{i,n}(x)$ only considers probability map of one ROI and (2) all the elements in $\theta_i(x)$ are equally treated, regardless of their distances to the center point x .

Predict ROI Labels in the Testing Stage—Given a sequence of learned classifiers, the conventional ACM iteratively predicts the classification probability for each point in the target image by gradually refining the probability map. Specifically, to label ROI O_n , both image appearance features and context features in the k -th iteration are extracted,

respectively, from the target image T and the classification probability map $P_{T,n}^{(k-1)}$ estimated in the $(k-1)$ -th iteration. Then, we apply the learned k -th classifier to these appearance/

context features for estimating the new classification probability map $P_{T,n}^{(k)}$ to update the context features for the $(k+1)$ -th iteration. We repeat these steps until all K classifiers are applied. In the end, we can segment the ROI O_n in the target image based on the final estimated classification probability map.

2.2. Multi-ROI ACMs for Joint Labeling of ROIs

Since the context feature in the conventional ACM is defined from a particular ROI O_n , it is unaware of the geometric relationship with neighboring ROIs. However, such cross-ROI information can be very useful to accurately label multiple neighboring ROIs. In this section, we extend the conventional ACM to handle multiple ROIs jointly.

Specifically, we will jointly train a set of ACMs for all spatially-connected ROIs, to allow them to share their respective intermediate outputs during the training. When training the ACM for the ROI O_n at the k -th iteration, we will not use only its own classification probability map $\{P_{i,n}^{(k-1)}, i = 1, \dots, M\}$, but the classification probability maps of all ROIs, $\{P_{i,n}^{(k-1)}, i = 1, \dots, M, n = 1, \dots, N\}$, to obtain the cross-ROI context features $\Theta_i(x) = [\theta_{i,1}(x), \dots, \theta_{i,n}(x), \dots, \theta_{i,N}(x)]$. It is clear that $\Theta_i(x)$ is the augment of all ROI-specific context features in all neighboring ROIs. Then, we can further train the ACM for each ROI O_n by using the obtained image appearance and cross-ROI context features. Different ACMs infer different sets of classifiers at the k -th iteration, which are used to generate the new classification probability maps for the next iteration (the $k+1$ -th iteration).

2.3. Improved ACM with Multi-Scale Representation and Multi-Atlases Segmentation Paradigm

Multi-Scale Representation for Probability Map—In conventional ACM, every element in the ROI-specific context feature $\theta_{i,n}(x)$ uses the probability value at a fixed scale. However, the displacement of each sampling location in δ has different importance for the center point. Therefore, it is not desirable to use the same scale to describe the shape priors at different sampling locations. Here, we allow each sampling location to use adaptive scale for capturing the shape priors from the classification probability map. Specifically, we partition the whole sampling pattern δ into several nested non-overlapping layers, spreading from the center point to the bound of the probability map. We then apply Gaussian smoothing to the probability map with different kernel sizes. Next, we use the smoothed probability map with the smallest kernel size to capture the fine-scale shape prior for the layer (of sampling locations) closest to the patch center. Gradually, we use the smoothed probability map with larger scale to capture the coarse-scale shape prior as the distance between the sampling location and the patch center increases. In this way, the image patch can be equipped with the multi-scale context feature representation.

Fig. 1(b) demonstrates the construction of multi-scale context feature $\tilde{\theta}_{i,n}(x)$. First, we smooth the original probability map $P_{i,n}$ with increasing Gaussian kernel sizes, to obtain multi-scale probability maps $\{P_{i,n,s} | s = 1, \dots, S\}$. Then, from the center point outward the boundary of probability map, we sequentially partition all sampled locations δ_q into S nested layers $\{Y_s | s = 1, \dots, S\}$. As displayed by the different sizes of blue boxes in Fig. 1(b), the smoothing kernel size become larger and larger outwards the center point. Finally, the multi-scale context feature $\tilde{\theta}_{i,n}(x) = [P_{i,n,s}(x + \delta_q) | q = 1, \dots, Q, \delta_q \in Y_s]$ can be computed. It is clear that $\tilde{\theta}_{i,n}(x)$ has the same vector length with $\theta_{i,n}(x)$ but use different scales at different sampling locations. After replacing $\theta_{i,n}(x)$ with $\tilde{\theta}_{i,n}(x)$, we can obtain cross-ROI and multi-scale context feature $\tilde{\Theta}_i(x)$.

Multi-Atlases Segmentation Paradigm—Multi-atlases segmentation paradigm can significantly improve the segmentation performance since the atlases encompass the high structure variations in the population. It is straightforward to simultaneously segment multiple ROIs by using multiple atlases. Given M training images I and their (ground-truth) manual labels L as mentioned above, we can train the cross-ROI and multi-scale ACM in the

domain of each training image I_i by registering all other training image I_j ($j \neq i$) to I_i . In the testing stage, we treat all training images as the atlas images. First, we sequentially register each training image I_i to the target image T . Thus, we can bring the learned cross-ROI and multi-scale ACMs, trained in each atlas domain, to the target image domain. Next, we allow each cross-ROI and multi-scale ACM to predict the probability maps for multiple ROIs. Thus, for each ROI under consideration, we have $M \times N$ classification probability maps, which encode the estimated segmentation information from multiple atlases for multiple ROIs. After we average all these probability maps based on the appearance discrepancies between target image and each aligned atlas image, we can use a level-set algorithm [8] to determine the final binary segmentation result for each ROI.

3. Experimental Results

In the following experiments, we comprehensively evaluate the performance of our *e*ACM labeling method on 40 images in LONI LPBA40 dataset, where each image has 56 manually labeled ROIs. 20 images are used as training images, while the other 20 images are used as testing images. To evaluate the performance of the proposed *e*ACM method for simultaneously labeling multiple ROIs, we specifically examine the segmentation results in two groups of spatially-connected ROIs, i.e., **ROISet1** (hippocampus, parahippocampus gyrus, fusiform gyrus), and **ROISet2** (caudate, putamen, insular cortex). Since the ROIs in both ROISet1 and ROISet2 are symmetrically defined in both left and right hemispheres, we apply our *e*ACM for each hemisphere separately. In both training stage and testing stage, the affine registration is used to map all images to the same reference space by using the FLIRT algorithm in FSL package. Both qualitative and quantitative evaluation results on the 20 testing images are provided in the next, by comparing with the conventional ACM method which also uses the multiple atlases. For fair comparison, we use the same parameters for the ACM used in both methods, e.g., the feature sets and the number of iterations.

3.1 Qualitative Evaluation on Labeling Results

Fig. 2 shows the labeling results for the three ROIs from ROISet1 (the 1st row of Fig. 2) and also another three ROIs from ROISet2 (the 2nd row of Fig. 2) on a selected test image by the conventional ACM method and our *e*ACM method. Compared with manual segmentations (Fig. 2 (a) and (d)), it is apparent that the labeling results by our method (Fig. 2 (c) and 2(f)) for both ROI sets are much closer to the manual segmentations than the conventional ACM method (Fig. 2 (b) and 2(e)), especially on the boundary areas between different ROIs.

3.2 Quantitative Evaluation on Labeling Results

For quantitative comparison of the conventional ACM and our *e*ACM, we use 4 overlap metrics (Precision, Recall, Jacquard, and Dice ratio) between the manual labeling and automatic labeling results by the conventional ACM and our *e*ACM, respectively. The

metrics are defined as $\text{Precision} = \frac{V(A \cap B)}{V(B)}$, $\text{Recall} = \frac{V(A \cap B)}{V(A)}$, $\text{Jacquard} = \frac{V(A \cap B)}{V(A \cup B)}$ and $\text{Dice} = \frac{V(A \cap B)}{\{V(A) + V(B)\}/2}$, where $V(A)$ is the volume of manual labeling and $V(B)$ is the

volume of automatic labeling result. Here, we specifically evaluate hippocampus (HP), parahippocampus gyrus (PG), and fusiform gyrus (FG) in ROISet1, and also caudate (CA), putamen (PU), and insular cortex (IC) in ROISet2, since they are all spatially-connected in each ROI set. Table 1 shows the average values (averaged over 20 testing images) of these four metrics in these 6 ROIs by the conventional ACM and our *e*ACM, respectively. It is clear our *e*ACM outperforms the conventional ACM in all evaluation metrics and all examined ROIs, indicating the advantages of our *e*ACM in simultaneously handling multiple ROIs and improving the labeling robustness by using multi-scale representations.

4. Conclusion

In this paper, we present an enhanced ACM (*e*ACM) method to simultaneously label a set of ROIs in the brain images. Specifically, we jointly train multiple cross-ROI ACMs for joint labeling of the neighboring ROIs, by allowing different ACMs to share their intermediate classification probability maps during the training and testing procedures. Thus, these ACMs can capture the geometric heuristics of all neighboring ROIs during the iterative labeling. We further improve the robustness of labeling by upgrading the probability maps with multi-scale representations and further integrating them to a multi-atlases paradigm. After we incorporate all improvements into the conventional ACM method, more accurate segmentation results are achieved by our *e*ACM method on the LONI LBPA 40 dataset.

References

1. Tu Z, Bai X. Auto-Context and Its Application to High-Level Vision Tasks and 3D Brain Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2010; 32:1744–1757. [PubMed: 20724753]
2. van der Lijn F, et al. Automated Brain Structure Segmentation Based on Atlas Registration and Appearance Models. *Medical Imaging, IEEE Transactions on*. 2012; 31:276–286.
3. Zhang S, et al. 3D Segmentation of Rodent Brain Structures Using Hierarchical Shape Priors and Deformable Models. *Medical Image Computing and Computer-Assisted Intervention – MICCAI*. 2011:611–618. [PubMed: 22003750]
4. Liu, X., et al. Deformable Atlas for Multi-structure Segmentation. In: Mori, K., et al., editors. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*. Vol. 8149. Springer; Berlin Heidelberg: 2013. p. 743-750.
5. Morra JH, et al. Comparison of AdaBoost and Support Vector Machines for Detecting Alzheimer's Disease Through Automated Hippocampal Segmentation. *Medical Imaging, IEEE Transactions on*. 2010; 29:30–43.
6. Powell S, et al. Registration and machine learning-based automated segmentation of subcortical and cerebellar brain structures. *NeuroImage*. 2008; 39:238–247. [PubMed: 17904870]
7. Coupé P, et al. Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. *NeuroImage*. 2011; 54:940–954. [PubMed: 20851199]
8. Chan TF, Vese LA. Active contours without edges. *Image Processing, IEEE Transactions on*. 2001; 10:266–277.

Fig. 1.
The extraction of context features in conventional ACM (a) and our eACM (b).

Fig. 2.

Visual comparison of labeled ROIs from 2 ROI sets. Compared to the manual labels (a) and (d), the labeling results (c) and (f) by our proposed *eACM* method are more accurate than the results (b) and (e) by the conventional ACM.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Quantitative comparisons between conventional ACM and our *e*ACM on 6 ROIs (hippocampus, parahippocampus gyrus, fusiform gyrus, caudate, putamen, and insular cortex), based on the Precision, Recall, Jacquard and Dice ratio measurements. (unit: %).

Table 1

	Metric	HP	PG	FG	CA	PU	IC
Conventional ACM	Precision	79.4	82.9	82.5	88.6	87.4	86.4
	Recall	80.2	83.1	83.6	90.0	88.2	85.9
	Jacquard	78.2	80.6	81.3	88.9	86.6	84.8
	Dice	80.6	82.9	82.8	89.8	88.9	86.0
Our <i>e</i> ACM	Precision	80.8	84.1	82.9	88.7	87.9	87.2
	Recall	81.1	85.4	84.2	90.2	88.4	86.3
	Jacquard	79.4	81.2	81.4	89.1	86.9	85.2
	Dice	81.4	83.8	83.1	90.1	89.2	86.8