Feature extraction using convolutional neural networks for multi-atlas based image segmentation

Xuesong Yang, Yong Fan

Xuesong Yang, Yong Fan, "Feature extraction using convolutional neural networks for multi-atlas based image segmentation," Proc. SPIE 10574, Medical Imaging 2018: Image Processing, 1057439 (2 March 2018); doi: 10.1117/12.2293876

SPIE.

Event: SPIE Medical Imaging, 2018, Houston, Texas, United States

Feature extraction using convolutional neural networks for multi-atlas based image segmentation

Xuesong Yang[#§], Yong Fan*
and for the Alzheimer's Disease Neuroimaging Initiative[Δ]

[#]National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academic of Sciences, Beijing 100190, China
[§]University of Chinese Academic of Sciences, Beijing 100190, China
[*]Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

ABSTRACT

Multi-atlas based image segmentation in conjunction with pattern recognition based label fusion strategies has achieved promising performance in a variety of image segmentation problems, including hippocampus segmentation from MR images. The pattern recognition based label fusion consists of image feature extraction and pattern recognition components. Since the feature extraction component plays an important role in the pattern recognition based label fusion, a variety of feature extraction methods have been proposed to extract image features, including texture features and random projection features. However, these feature extraction methods are not adaptive to different segmentation problems. Following the success of convolutional neural networks in image feature extraction, we propose a feature extraction method based on convolutional neural networks for multi-atlas based image segmentation. The proposed method has been validated based on 135 T1 magnetic resonance imaging (MRI) scans and their hippocampus labels provided by the EADC-ADNI harmonized segmentation protocol. We also compared our method with state-of-the-art pattern recognition based MAIS methods, including Local Label Learning and Random Local Binary Patterns. The experimental results have demonstrated that our method could achieve competitive hippocampus segmentation performance over the alternative methods under comparison.

Keywords: Multi-atlas based image segmentation, hippocampus, feature extraction, convolutional neural networks

1. INTRODUCTION

Medical image segmentation is an important component for delineating regions of interest in medical image analysis. For instance, delineation of the hippocampus from magnetic resonance imaging (MRI) scans has received much attention in studies of Alzheimer's disease [1]. Multi-atlas based image segmentation (MAIS) is one of the most successful strategies for the hippocampus segmentation problem [2, 3]. The MAIS methods segment images based on a set of atlases consisting of different images and their corresponding segmentation label images in following 2 steps. First, all the atlas images including their corresponding segmentation label images are registered to the image to be segmented using image registration algorithms, and then a segmentation label fusion strategy is adopted to fuse the registered atlas label images to generate a segmentation result. Particularly, the label fusion strategy plays an important role in the MAIS methods.

A variety of label fusion strategies have been developed and most of them can be categorized as voting based methods, image representation based methods, and pattern recognition based methods. The voting based methods fuse the registered atlas label images using voting strategies, such as majority voting [4] and simultaneous truth and performance level estimation (STAPLE) [5]. The image representation based methods fuse the registered atlas label images by learning an image representation to characterize the image to be segmented based on the registered atlas images and applying the learned representation to the registered atlas label images to generate a segmentation label

[Δ] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Medical Imaging 2018: Image Processing, edited by Elsa D. Angelini, Bennett A. Landman, Proc. of SPIE
Vol. 10574, 1057439 · © 2018 SPIE · CCC code: 1605-7422/18/$18 · doi: 10.1117/12.2293876

image. The image representation can be learned using methods, such as sparse representation [6] and dictionary learning [7]. The pattern recognition based methods build voxel-wise classifiers by using the registered atlas images and label images as training data and apply the classifiers to the image to be segmented to generate a segmentation label image [2, 3, 8-13]. A variety of pattern recognition methods have been used to build such classifiers, including support vector machines [2, 3], random forests [14], and linear regression [10].

In the MAIS methods that are built upon pattern recognition based label fusion strategies, image feature extraction used in the classification plays an important role. In particular, most existing pattern recognition based MAIS methods use nonlocal patch-based (NLP) strategy [15] to generate local training samples for each voxel of images to be segmented. To build local classifiers for segmenting images, a variety of feature extraction techniques have been proposed to extract image features from image patches. Besides using original image intensity values as image features, the most frequently used image features are texture image features, such as those extracted using first and second order difference filters, Hyperplane filters, Sobel filters, Laplacian filters and range difference filters [2, 3], Haar features [16], and random local binary patterns [10]. All these feature extraction methods extract image features independent on the segmentation problems under study, and therefore they are not necessarily the best for the image segmentation problem under study.

Deep learning techniques, particularly convolutional neural networks (CNNs), can adaptively learn image features for a specific image segmentation problem under study [17]. Particularly, a deep CNN model is an integrated model of both feature generators and classifiers, and it could adjust weights of CNNs to optimize image feature extraction and classification. By integrating multiple convolutional layers in a single deep learning model, hierarchical image features can be extracted to capture spatial contextual information of images and therefore improve discriminative power of the images features.

In this study, we propose a pattern classification based MAIS method to segment the hippocampus from MRI scans by building voxel-wise classifiers upon image features extracted by CNNs. Our method first trains a global CNN model to extract image features from image patches and then builds voxel-wise classifiers based on nonlocal image patches of atlas images as training data. The proposed method has been validated based on 135 T1 MRI scans and their hippocampus labels provided by the EADC–ADNI (European Alzheimer's Disease Consortium and Alzheimer's Disease Neuroimaging Initiative) harmonized segmentation protocol (www.hippocampal-protocol.net) [18]. We also compared our method with state-of-the-art pattern recognition MAIS methods, including local label learning (LLL) [2] and RLBP [10]. The experimental results have demonstrated that our method could achieve competitive hippocampus segmentation performance over the alternative methods under comparison.

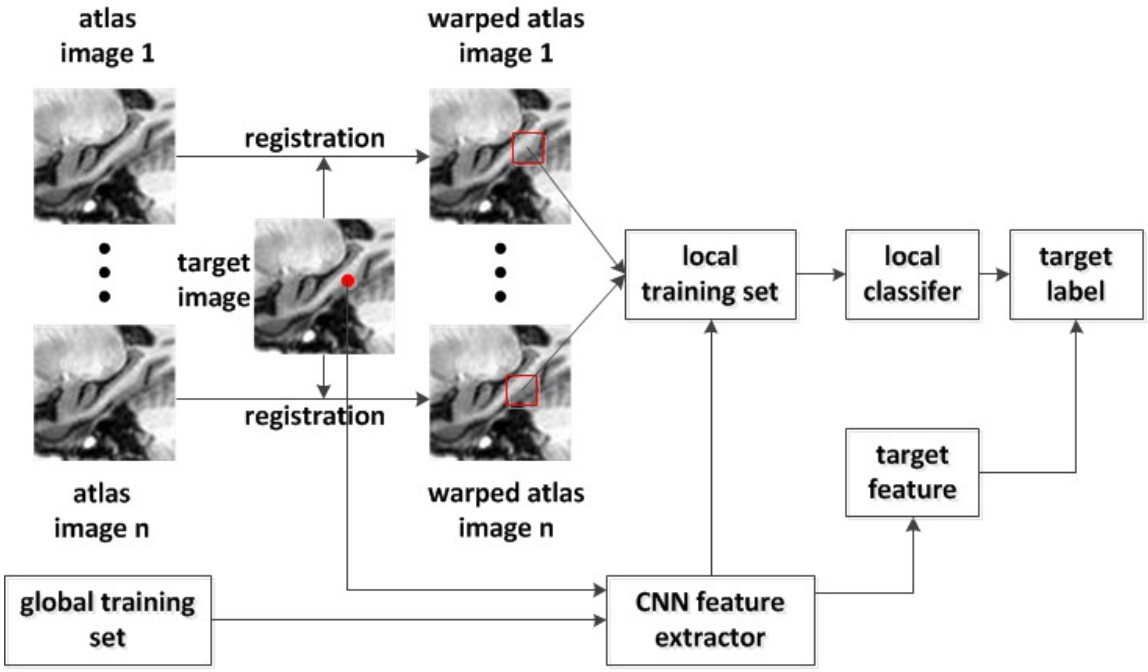

Figure 1. Flowchart of the multi-atlas based image segmentation framework.

2. METHODS

2.1 Multi-atlas based image segmentation

Given a target image I to be segmented and N pairs of atlases images and their corresponding segmentation label masks that have been already registered to the target image, $A_i = (I_i, L_i), i = 1,2, \dots, N,$ where I_i is the i-th atlas image and L_i is its segmentation label, a local classifier for each voxel p of the target image is to be built upon its neighboring voxels of the atlas images and their segmentation label images. Particularly, as shown in Figure 1, for each voxel $p_{i,j,k}$ of an image to be segmented, where (i, j, k) is its 3D coordinates, we first build a training dataset by collecting each atlas image's voxels that are located within a cubic neighborhood of (i, j, k) with a neighborhood size of $(2r_s + 1)^3$, and each of these voxels, including the voxel $p_{i,j,k}$ to be segmented, is characterized by a cubic image patch with a radius r_p. Then, we build a CNN based deep learning model to extract CNN features and build classifiers to classify the image patches into foreground and background classes based on the training data. Finally, the learned CNN feature extractors are used to extract image features for building local classifiers to segment images in a MAIS framework. It is worth noting that the CNN based deep learning model is built upon a separate set of atlases and used to extract image features for the pattern recognition based MAIS method.

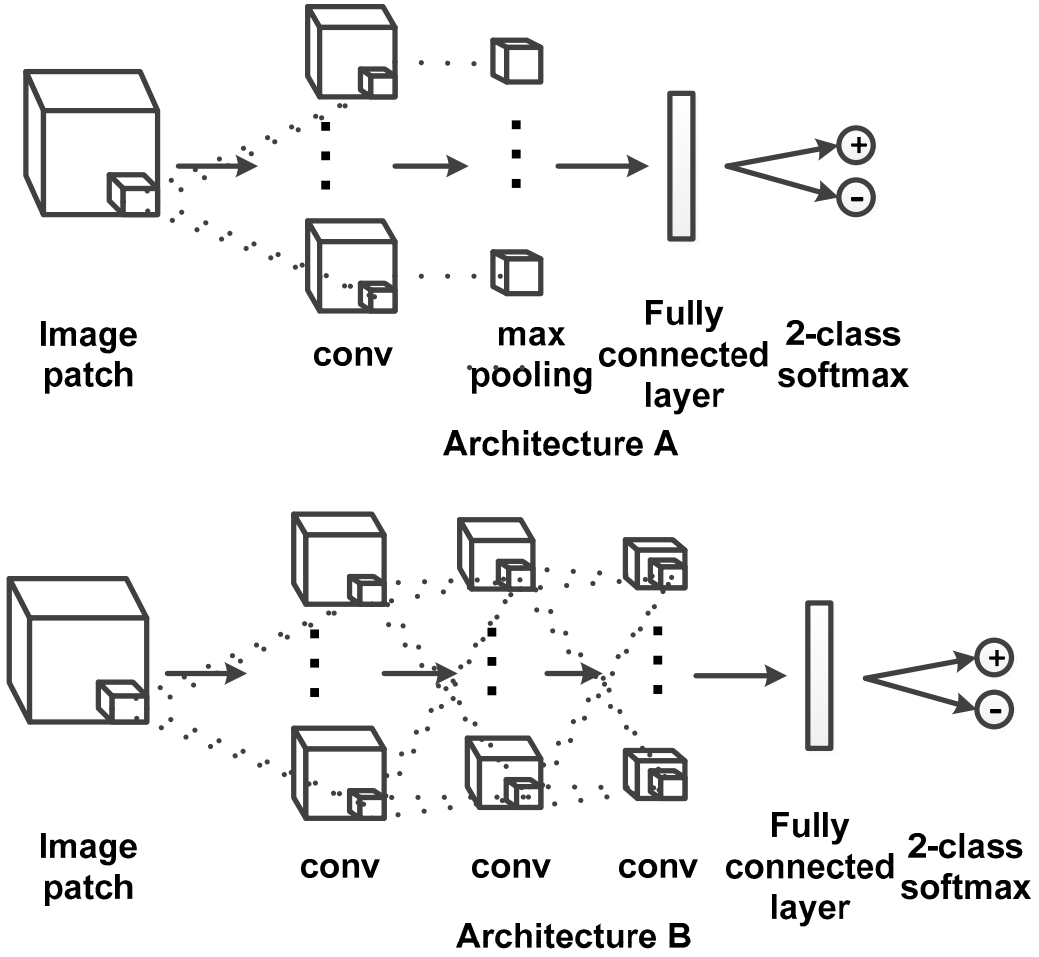

Figure. 2. Flowchart of the CNN based image feature extractors. Convolution layer is abbreviated as conv.

2.2 Feature extraction using CNNs

In this study, we build a CNN based deep learning model to extract image features within the MAIS framework. Particularly, we explore 2 different architectures of CNNs to extract image features, as shown in Figure 2. Specifically, the first network architecture, referred to as architecture A, consists of one convolutional layer, one max pooling layer, one fully connected layer, and one two-class softmax classification layer. The second network architecture, referred to as architecture B, consists of three convolutional layers, one fully connected layer, and one two-class softmax layer. There are 3 parameters in each of network architectures, including size of input image patches, kernel size of the convolutional layers, and the number of kernels. The max pooling layer's kernel size and stride are set to $2 \times 2 \times 2$ and 2 respectively for all the models. The number of nodes of all the fully connected layers is set to 512. Activation functions of all convolutional layers are implemented by rectified linear units (ReLUs). The convolution kernels' weights and biases are initialized based on uniform distribution U(-1,1) and are updated to optimize image features guided by a softmax classifier with a cross-entropy loss [19]. The input image patches of all networks are preprocessed by taking the following steps: 1) normalizing voxels of each image patch by subtracting their mean value and dividing by their standard deviation; and 2) subtracting value of center voxel from all voxels of the image patch.

To train the deep learning model with different architectures, we construct a training set with image patches randomly sampled from the training images. Once the deep learning models are trained, they are used to extract image features from image patches and output values of the max pooling layer in architecture A or the third convolutional layer in architecture B are used as image features, referred to as CNN features, for the MAIS based image segmentation. Once the CNN features are extracted for image patches, we build a local classifier on them to predict segmentation label of each voxel of images to be segmented using the pattern recognition based label fusion strategy [2, 3].

2.3 Voxel-wise image classification

Following the pattern recognition based label fusion strategy [2, 3], we build a local classifier for each voxel of images to be segmented based on non-local image patches of the atlas images [15]. Therefore, for each voxel p to be segmented, we first identify its neighboring voxels in the atlas images and segmentation label images from atlas images to build a training dataset, then extract CNN features from image patches of all image voxels of the training dataset and the voxel to be segmented, and finally build a linear regression model to segment the voxel under consideration based on the CNN features. Given a local training dataset of voxel p with n training samples, $D_p = \{(f_{i,p}, L_{i,p}), i = 1, \dots, n\}$, where $f_{i,p}$ and $l_{i,p}$ are CNN features and segmentation label of the i-th training sample respectively, a linear regression model is built by optimizing

$$\beta = \arg\min \frac{1}{2}\|\beta\|_2^2 + C\frac{1}{2}\sum\left(l_{i,p} - f_{i,p}\beta\right)^2, \quad (1)$$

where β is a coefficient vector to be learned, and C is a regularization coefficient.

Once the linear regression model is built, voxel p's segmentation label l_p can be inferred directly by thresholding the regression model's output

$$l_p = \begin{cases} 1, f_p\beta \geq 0.5 \\ 0, f_p\beta < 0.5 \end{cases}, \quad (2)$$

where f_p is CNN features of the voxel p and β is the learned linear regression model's coefficient vector.

3. EXPERIMENTAL RESULTS

3.1 Imaging data

Two sets of MRI scans with manual hippocampus segmentation labels were used in our study. Particularly, the first set of MRI scans and their hippocampus labels were obtained from our previous study [2], including 57 3T MRI scans of different subjects obtained from a local hospital and 1.5T and 3T MRI scans of 30 subjects obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu/). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). The ADNI MRI scans were acquired using a sagittal 3D MP-RAGE T1-w sequence (TR = 2400 ms, minimum full TE, TI = 1000 ms, FOV = 240 mm, voxel size of $1.25 \times 1.25 \times 1.2\text{mm}^3$) [20]. For up-to-date information of the ADNI project, see www.adni-info.org. The second set of MRI scans and their hippocampus labels were obtained from the EADC-ADNI harmonized segmentation protocol, including MRI

scans of a preliminary release with 100 subjects and a final release with 35 subjects [18]. All the MRI scans were preprocessed using a standard protocol, including bias field correction and spatial normalization to the standard Montreal Neurological Institute (MNI) space using affine image registration [2].

The first set of MRI scans was used to train the CNN based deep learning models and tune parameters of the pattern recognition based MAIS method. Particularly, 95360 image patches and their corresponding hippocampus segmentation labels were randomly generated based on MRI scans of the first set and the final release of the second set. A validation dataset, consisting of 5 1.5T MR images and 5 3T MR images randomly selected from the first set of MR scans, was used to tune parameters of the pattern recognition based MAIS method to optimize its segmentation performance. The atlases used in the MAIS method were other MRI scans of the first set.

The segmentation performance of our method was finally evaluated based on the second set of MRI scans. Particularly, the MRI scans of the preliminary release part were used as testing data and the MRI scans of the final release part were used as atlas images. We also compared our method with 2 state-of-the-art pattern recognition based hippocampus segmentation methods, namely LLL and RLBP [2, 3, 10], based on the same testing data and atlas images.

To reduce computational cost of the MAIS methods under comparison, bounding boxes of bilateral hippocampi were determined using the same procedure used in our previous study [2], and the MAIS methods were applied to MRI scans within the bounding boxes for segmenting the hippocampi. To segment each testing MRI scan, top 20 most similar atlas images were selected based on max mutual information between the atlas images and the testing MRI scan within the bounding box [2], and the top 20 most similar atlas images were registered to the testing MRI scan using a non-grid image registration algorithm, namely ANTs [21]. The same strategy was used in all the MAIS segmentation experiments, including the experiment for tuning parameters of our method based on the validation dataset.

We adopted 3 metrics to evaluate the automatic segmentation results, including Dice coefficient, Mean Distance (MD), and Average Symmetric Surface Distance (ASSD), defined as

$$\mathrm{Dice} = 2\frac{\mathrm{V(A \cap B)}}{\mathrm{V(A) + V(B)}},$$

$$\mathrm{MD} = \mathrm{mean}_{e\in\partial A}\left(min_{f\in\partial B} d(e,f)\right),$$

$$\mathrm{ASSD} = \frac{\left(\mathrm{mean}_{e\in\partial A}\left(min_{f\in\partial B} d(e,f)\right) + \mathrm{mean}_{e\in\partial B}\left(min_{f\in\partial A} d(e,f)\right)\right)}{2},$$

where A is the manual segmentation result, B is the automatic segmentation result, $V(X)$ is the volume of segmentation X, ∂X as boundary voxels of X, and $d(\cdot,\cdot)$ is the Euclidian distance.

3.2 Parameter settings of the deep learning models

Table 1 summarizes parameters of the deep learning models used in our experiments. Specifically, the deep learning model of network architecture A was trained with 2 different settings for the image patch size: 7 and 9; 3 different settings for the number of convolutional kernels: 128, 256, and 512; and 2 different settings of the kernel size: 3 and 5. The deep learning model of network architecture B was trained with the image patch size set to 7, the number of convolutional kernels set to 128, and the kernel size set to 3. In total, there were 6 CNN based deep learning models were trained based on the training image atlases.

Table 1. Parameters setting of the deep learning models

CNN models	Network architecture	Image patch size	Number of kernels	Size of kernels
Model 1	A	7	128	3
Model 2	A	7	256	3
Model 3	A	7	512	3
Model 4	A	9	128	3
Model 5	A	9	128	5
Model 6	B	7	128	3

3.3 Parameter settings of the linear regression model used in the MAIS method

The linear regression model's parameter C was tuned based on the validation dataset for image features extracted by different CNN deep learning models based on the hippocampus segmentation performance measured using Dice index. As summarized in Table 2, the optimal parameter C varied for image features extracted by the different deep learning models.

Table 2. Average Dice index values of the bilateral hippocampi of MRI scans of the validation data set and optimal values of C.

C	1e-4	1e-3	1e-2	Optimal values
Model 1	0.9163	**0.9182**	0.9080	0.001
Model 2	**0.9186**	0.9183	0.9057	0.0001
Model 3	**0.9182**	0.9151	0.9034	0.0001
Model 4	**0.9185**	0.9173	0.9055	0.0001
Model 5	**0.9190**	0.9137	0.8927	0.0001
Model 6	0.9018	**0.9176**	0.9140	0.001

3.4 Implementation details of the CNN deep learning models

The CNN based deep learning models were implemented based on Tensorflow[1] and MexConv3D[2]. The models were trained using Adam optimizer [22] with a learning rate of 0.0001. Specifically, no padding operation was applied to any convolutional layers and max pooling layer (padding = 'VALID' in Tensorflow code), and strides for all kernels of convolutional layers were set to 1. The models were trained for 750000 epochs based on 95360 training image patches, and for each epoch the models were trained with a batch size equal to 128. It took ~2 hours to train the model on a Tesla K80c GPU.

3.5 Segmentation performance of the MAIS methods

The segmentation performance of our method with different parameter settings and alternative methods under comparison was estimated based on the testing dataset, as summarized in Table 3. All the MAIS methods build upon image features extracted by different deep learning models achieved promising segmentation performance. Particularly, the deep learning model 6 with network architecture B achieved the overall best segmentation performance, indicating that the image features extracted with hierarchical CNNs could better capture spatial contextual information of images. Results of Wilcoxon signed-rank tests also indicated that model 6 had significantly better performance than LLL, RLBP, and other deep learning models in terms of Dice index and Jaccard index. However, the differences between model 6 and some of the alternative methods in terms of MD and ASSD values were moderate. Sample segmentation results obtained by different methods are shown in Figure 3.

Table 3. Segmentation performance of all the methods under comparison.

Methods	Dice (left/right)	Jaccard(left/right)	MD (left/right)	ASSD (left/right)
LLL	0.8695/0.8771	0.7701/0.7820	0.3515/0.3233	0.3379/0.2885
p-value	*4.1e-17/1.3e-10*	*4.1e-17/8.7e-11*	*3.9e-18/2.6e-17*	*1.7e-15/4.0e-10*
RLBP	0.8749/0.8787	0.7784/0.7844	0.3080/0.2914	0.2804/0.2755
p-value	*2.0-07/3.6e-08*	*1.8e-07/3.3e-08*	*0.497/0.249*	*0.142/0.324*
Model 1	0.8766/0.8815	0.7811/0.7889	0.3081/**0.2887**	0.2898/0.2783
p-value	*0.012/0.411*	*0.011/0.374*	*0.942/0.293*	*0.003/0.942*
Model 2	0.8761/0.8813	0.7804/0.7887	0.3044/0.2850	0.2889/0.2773
p-value	*5.9e-07/0.001*	*5.5e-07/9.1e-04*	*5.6e-04/1.0e-05*	*0.038/0.827*
Model 3	0.8746/0.8795	0.7779/0.7856	0.3014/0.2841	0.2902/0.2813
p-value	*5.0e-11/2.9e-10*	*2.9e-11/2.4e-10*	*3.70e-05/2.2e-04*	*0.005/7.3e-04*
Model 4	0.8759/0.8808	0.7799/0.7877	0.3080/0.2911	0.2909/0.2761
p-value	*0.001/0.012*	*7.6e-04/0.010*	*0.858/0.310*	*0.064/0.215*
Model 5	0.8738/0.8780	0.7766/0.7833	0.3023/0.2871	0.2914/0.2789
p-value	*6.0e-08/5.0e-08*	*4.6e-08/4.1e-08*	*0.022/0.364*	*0.016/ 0.011*
Model 6	**0.8778/0.8816**	**0.7830/0.7892**	**0.3078**/0.2910	**0.2800/0.2758**

4. CONCLUSIONS

In this study, we demonstrated that image features extracted using CNN based deep learning models could improve image segmentation performance of the pattern recognition based MAIS methods. In particularly, our experimental results indicated that the CNN based deep learning models with multiple convolutional layers could extract image features with improved discriminative power than those extracted by the deep learning models with one convolutional

[1] http://tensorflow.org

[2] https://github.com/pengsun/MexConv3D

layer. Our ongoing study is devoted to the development of a deep learning based on hippocampus segmentation method by extending our brain tumor segmentation method [23].

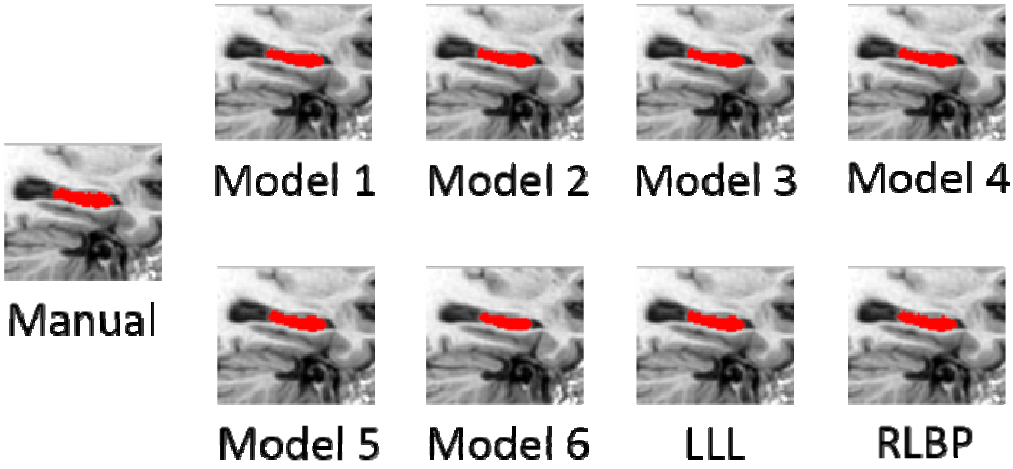

Figure 3. Sample segmentation results obtained by different methods under comparison.

ACKNOWLEDGEMENTS

This work was supported in part by the National Key Basic Research and Development Program of China [2015CB856404]; the National Natural Science Foundation of China [61473296].

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

REFERENCES

[1] R. Wolz, A. J. Schwarz, P. Yu *et al.*, "Robustness of automated hippocampal volumetry across magnetic resonance field strengths and repeat images," Alzheimers & Dementia, 10(4), 430-438 (2014).

[2] Y. Hao, T. Wang, X. Zhang *et al.*, "Local label learning (LLL) for subcortical structure segmentation: application to hippocampus segmentation," Hum Brain Mapp, 35(6), 2674-97 (2014).

[3] Y. Hao, J. Liu, Y. Duan *et al.*, "Local label learning (L3) for multi-atlas based segmentation," SPIE Medical Imaging, 83142E, 83142E (2012).

[4] T. Rohlfing, R. R. Brandt, and M. C. Jr, "Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains," Neuroimage, 21(4), 1428-1442 (2004).

[5] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation," IEEE Transactions on Medical Imaging, 23(7), 903-921 (2004).

[6] Y. Wu, G. Liu, M. Huang *et al.*, "Prostate Segmentation Based on Variant Scale Patch and Local Independent Projection," IEEE Transactions on Medical Imaging, 33(6), 1290-303 (2014).
[7] T. Tong, R. Wolz, Z. Wang *et al.*, "Discriminative dictionary learning for abdominal multi-organ segmentation," Medical Image Analysis, 23(1), 92-104 (2015).
[8] H. Zhu, H. Cheng, X. Yang *et al.*, "Metric learning for multi-atlas based segmentation of hippocampus," Neuroinformatics, 15(1), 41-50 (2017).
[9] H. Zhu, H. Cheng, X. Yang *et al.*, "Metric learning for label fusion in multi-atlas based image segmentation," 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), 1338-1341 (2016).
[10] H. Zhu, H. Cheng, and Y. Fan, "Random local binary pattern based label learning for multi-atlas segmentation," SPIE Medical Imaging, 94131B-94131B-8 (2015).
[11] Q. Zheng, and Y. Fan, "Integrating semi-supervised label propagation and random forests for multi-atlas based hippocampus segmentation," The IEEE International Symposium on Biomedical Imaging (ISBI), (2018).
[12] Y. Hao, T. Jiang, and Y. Fan, "Shape-constrained multi-atlas based segmentation with multichannel registration," Proceeding of SPIE Medical Imaging: Image Processing, 8314, 83143N (2012).
[13] Y. Hao, T. Jiang, and Y. Fan, "Iterative multi-atlas based segmentation with multi-channel image registration and Jackknife Context Model," 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), 900-903 (2012).
[14] Y. Shao, Y. Guo, Y. Gao *et al.*, "Hippocampus Segmentation from MR Infant Brain Images via Boundary Regression," International MICCAI Workshop on Medical Computer Vision, 9601, 146-154 (2016).
[15] P. Coupé, J. V. Manjón, V. Fonov *et al.*, "Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation," Neuroimage, 54(2), 940-54 (2011).
[16] L. Zhang, Q. Wang, Y. Gao *et al.*, "Concatenated Spatially-localized Random Forests for Hippocampus Labeling in Adult and Infant MR Brain Images," Neurocomputing, 229, 3-12 (2016).
[17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, 521, 436 (2015).
[18] M. Boccardi, M. Bocchetta, F. C. Morency *et al.*, "Training labels for hippocampal segmentation based on the EADC-ADNI harmonized hippocampal protocol," Alzheimers & Dementia, 11(2), 175 (2015).
[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems, 2012 (2012).
[20] C. R. Jack, M. A. Bernstein, N. C. Fox *et al.*, "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," Journal of Magnetic Resonance Imaging, 27(4), 685-691 (2008).
[21] B. B. Avants, C. L. Epstein, M. Grossman *et al.*, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," Medical Image Analysis, 12(1), 26–41 (2008).
[22] D. Kingma, and J. Ba, "Adam: A Method for Stochastic Optimization," arXiv:1412.6980, (2014).
[23] X. Zhao, Y. Wu, G. Song *et al.*, "A deep learning model integrating FCNNs and CRFs for brain tumor segmentation," Med Image Anal, 43, 98-111 (2018).