Multi-Atlas Based Methods in Brain MR Image Segmentation

Liang Sun¹, Li Zhang², Daoqiang Zhang¹*

¹College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, MIIT Key Laboratory of Pattern Analysis and Machine Intelligence, Nanjing 211106, China
²Department of Geriatrics, the Affiliated Brain Hospital of Nanjing Medical University, Nanjing 210029, China

Key words: multi-atlas; brain; segmentation; magnetic resonance

Abstract
Brain region-of-interest (ROI) segmentation is an important prerequisite step for many computer-aided brain disease analyses. However, the human brain has the complicated anatomical structure. Meanwhile, the brain MR images often suffer from the low intensity contrast around the boundary of ROIs, large inter-subject variance and large inner-subject variance. To address these issues, many multi-atlas based segmentation methods are proposed for brain ROI segmentation in the last decade. In this paper, multi-atlas based methods for brain MR image segmentation were reviewed regarding several registration toolboxes which are widely used in the multi-atlas methods, conventional methods for label fusion, datasets that have been used for evaluating the multi-atlas methods, as well as the applications of multi-atlas based segmentation in clinical researches. We propose that incorporating the anatomical prior into the end-to-end deep learning architectures for brain ROI segmentation is an important direction in the future.
Since MR images often suffer from low intensity contrast around the boundary of ROIs, large inter-subject variance and large inner-subject variance, as well as human brain has complicated anatomical structure, ROI segmentation with the brain MR image is a difficult task for conventional machine learning methods. To solve these issues, multi-atlas based methods are proposed for brain ROI segmentation by using the anatomical prior of brain. The basic idea of multi-atlas is that the voxels should share the same label if they have similar local patterns or appearances. According to the basic idea, the multi-atlas based methods typically include two steps, 1) image registration, and 2) label fusion. In the step of registration, the atlas images are registered onto the common space of to-be-segmented image; in the step of label fusion, the label of voxels in atlases are propagated to the to-be-segmented image to determine the final label.

In this paper, we first present several registration toolboxes which are widely used in the multi-atlas methods, and describe some conventional methods for label fusion. Also, we introduce some datasets for evaluating the multi-atlas methods, as well as the applications of multi-atlas based segmentation in clinical researches. Finally, a brief conclusion is given in the Section Conclusion.

REGISTRATION

In multi-atlas based segmentation methods, the label fusion is based on the similarity of local pattern or appearance between the voxels in the to-be-segmented images and the candidate voxels in the atlas images. Hence, the segmentation performance of multi-atlas method relies on the image alignment results. In multi-atlas based segmentation method, the atlas images and to-be-segmented image are firstly registered onto the common space. The registration methods commonly include the affine registration and deformed registration. In this section, we present several widely used registration toolboxes in literature for multi-atlas based segmentation with brain MR images.

FMRIB’s Linear Image Registration Tool

FMRIB’s Linear Image Registration Tool (FLIRT)\(^6\)\(^,\)\(^7\) in the FMRIB Software Library (FSL)\(^a\)\(^,\)\(^b\) is an affine registration method, which provides a standard framework for MR images registration based on the image intensity. FLIRT can handle multi-modality medical image (e.g., functional MRI, structure MRI and PET). FLIRT uses the global optimization strategy for modifying the cost functions to down-weight voxels at the edge of the common overlapping field-of-view, and fuzzy-binning techniques for histogram estimation to reduce the local minimization value of the cost function. Hence, FLIRT is more efficient for searching the parameters, and further speeds the registration process. FLRIT provides many cost functions for optimization, such as normalized correlation, correlation ratio, mutual information, normalized mutual information and least square, etc.

ANTS Symmetric Normalization algorithm

ANTS Symmetric Normalization (SyN) algorithm\(^9\) in ANTs Normalization Tools\(^b\) is a deformed registration method, which developed for the subject with the brain disease, such as frontotemporal dementia (FTD) and Alzheimer’s disease (AD). Hence, the ANTS Syn could process not only the normal subjects, but also images with large distance from the template images. The diffeomorphism is a differentiable map with differentiable inverse. The ANTS Syn restricts the proposed method into the diffeomorphic with homogeneous boundary conditions and assume the rigid and scaling transformations are factored out and the image border maps to itself. The proposed method uses the geodesic distance to obtain the shortest paths between subjects in the diffeomorphic space. The method aims to find a spatiotemporal mapping to maximize the cross-correlation between the pairwise images in the space of diffeomorphic, then use the Euler-Lagrange equations necessary to optimize the cost function.

Efficient non-parametric image registration method

Vercauteren et al. proposed an efficient non-parametric image registration method\(^c\) based on the Thirion’s demons algorithm.\(^11\) The non-parametric image registration can be seen as an optimization problem, which aims to find the displacement of each voxel to obtain an alignment of images. Vercauteren et al. introduce the diffeomorphisms into the demons algorithm to preserve the topology of the objects in the image and prohibit the physically impossible folding. Due to the diffeomorphic image registration method often suffers

\(^6\) FMRIB's Linear Image Registration Tool (FLIRT) is an affine registration method, which provides a standard framework for MR images registration based on the image intensity. FLIRT can handle multi-modality medical image (e.g., functional MRI, structure MRI and PET).

\(^7\) FMRIB's Linear Image Registration Tool (FLIRT) is an affine registration method, which provides a standard framework for MR images registration based on the image intensity. FLIRT can handle multi-modality medical image (e.g., functional MRI, structure MRI and PET).

\(^8\) FMRIB Software Library (FSL) is an affine registration method, which provides a standard framework for MR images registration based on the image intensity. FLIRT can handle multi-modality medical image (e.g., functional MRI, structure MRI and PET).

\(^9\) ANTS Symmetric Normalization (SyN) algorithm is a deformed registration method, which developed for the subject with the brain disease, such as frontotemporal dementia (FTD) and Alzheimer’s disease (AD).

\(^10\) ANTs Normalization Tools is a deformed registration method, which developed for the subject with the brain disease, such as frontotemporal dementia (FTD) and Alzheimer’s disease (AD).

\(^11\) Thirion’s demons algorithm is a deformation method that uses the geodesic distance to obtain the shortest paths between subjects in the diffeomorphic space. The method aims to find a spatiotemporal mapping to maximize the cross-correlation between the pairwise images in the space of diffeomorphic, then use the Euler-Lagrange equations necessary to optimize the cost function.
from the computational burden, the authors also intro-
duced a non-parametric diffeomorphic method into the
demons algorithm. Vercauteren et al. used the Newton
methods for Lie groups to optimize the process of dif-
femorphic image registration.

LABEL FUSION

The aim of multi-atlas based segmentation meth-
method is to segment the target images according to the
anatomical priors from a set of atlas images. In last
decade, many label fusion strategies are proposed
for propagating the labels from atlases to determine
the final label of the to-be-segmented image. In this
section, we introduce several classic label fusion meth-
ods for segmenting the ROIs with brain images.

Major voting

Intuitively, the major voting (MV) strategy can be
used for determining the label of to-be-segmented
voxel. Specifically, the MV adopts the label with the
greatest number of labels at the same location of the
to-be-segmented voxel in the atlases as the final label
of the to-be-segmented voxel.

Locally-weight voting

As the simplest strategy, MV does not consider the
similarity of the local appearance or pattern between
the to-be-segmented voxel and the voxel in the atlas
image. Based on the assumption that the voxels
should share the same labels if the voxels have the
similar local appearance or pattern, locally-weighted
voting (LWV) method propagates the labels on the
atlases images at the same location of to-be-segmented
by the weight sum strategy. Specifically, the similarity
between the to-be-segmented voxel $y$ and the voxel $x_i$
in the atlas image $A_i$ can be calculated as follows,

$$ w(y, x_i) = \exp \frac{\| P(y) - P(x_i) \|^2}{\sigma} $$

where $P(y)$ and $P(x_i)$ are the vectorized intensity of
local image patch, which sit at the center of voxel $y$ and
voxel $x_i$. Also, the term $\sigma = \min_x \| P(y) - P(x_i) \| + \epsilon$, where $\epsilon$ is a small constant to ensure numerical stability. Then, the LWV uses the calculated similarity as the voting weight to calculate the soft label of to-be-segmented by the weight sum,

$$ l(y) = \arg \max_c \sum_{i=1}^N w(y, x_i) \delta(l(x_i) = c) $$

where $\delta(l(x_i) = c)$ is a Dirac function, which equals to 1
when $l(x_i) = c$ and 0 otherwise. Hence, patches that are
more similar to the target patch have greater voting
weight. Finally, we use the maximum a posteriori (MAP)
criterion to obtain the label of $y$,

$$ l(y) = \arg \max_c l(y = c) $$

Joint label fusion

Base on the Eq. (3), the similar patches on the atlases
often have the similar voting weights, which increase
the risk of the different atlases produce the similar la-
bel errors. To address this issue, the joint label fusion
(JLF) method is proposed for reducing the expecta-
tion of labeling errors between the pairwise atlases,

$$ w^* = \arg \max_y \sum_{i=1}^N w_i M_y w_i $$

$$ s.t. \sum_{i=1}^N w_i = 1 $$

where $w_i = \{w(y, x_i)\}_{i=1}^N$. $M_y$ is a matrix of expect-
ed pairwise joint label differences between atlas imag-
es and to-be-segmented image. In literature, the ele-
ment $M_y(i, j)$ in $M_y$ can be calculated by the local image
intensity information between two images to predict
their label difference,

$$ M_y(i, j) = \sum_{v \in A_i} | F(x) - F_A(v) | $$

where $v$ is a voxel belonging to the local region of the
to-be-segmented voxel $y$. $F(v)$ is the intensity of the
voxel. Hence, if both patches in the atlases have large
intensity differences from the patch in the target im-
age, the two patches are expected to produce different
labels. Finally, use the learnt voting weight $w_i$ for label
fusion by Eq. (2) and Eq. (3).

Non-local mean patch-based method

Although the LWV and JLF consider the local pattern
similarity between the to-be-segmented voxel and the
voxel in the atlas image, JLF also considers the differ-
ence between atlas images.

The registration step often leads to the registration
errors. To alleviate the possible registration errors, the
non-local mean patch-based method (PBM) is pro-
posed for label fusion. As shown in Figure 1, the PBM
propagates the labels not only from the voxel sitting
in the same location of the to-be-segmented images
in atlases, but also from its neighboring locations in
the atlases. Specifically, PBM seeks multiple candidate
patches from a certain region in each atlas for label fu-
sion based on the pairwise similarity between the target
and candidate patches within the certain region via
\[ w(y, x_{ik}) = \exp \left( \frac{-||(y-x_{ik})||^2}{\sigma} \right) \]  

(6)

where \( x_{ik} \) is the \( k \)-th voxel in \( y \)'s neighboring region \( N(y) \) in the \( i \)-th atlas image. Also, the term \( \sigma = \min_k ||P(y) - P(x_{ik})|| + \epsilon \). PBM uses the calculated pairwise similarity between the target patch and the selected candidate patch within the certain region as the voting weight to calculate the soft label of to-be-segmented by the weight sum,

\[ l(y = c) = \frac{\sum_{i=1}^{N} \sum_{k=1}^{K} w(y, x_{ik}) \delta(l(x_{ik}) = c)}{\sum_{i=1}^{N} \sum_{k=1}^{K} w(y, x_{ik})} \]  

(7)

where \( \delta(l(x_{ik}) = c) \) is a Dirac function, which equals to 1 when \( l(x_{ik}) = c \) and 0 otherwise. Also, PBM obtain the final label of to-be-segmented voxel by Eq. (3).

Due to the non-local strategy, the PBM can select the candidate patch in the atlases within the local regions of to-be-segmented voxel, and is more robust to the possible registration errors than LWV. Based on the non-local strategy, JLF also has been extended with non-local strategy. Instead of only selecting the candidates at the same location of the to-be-segmented voxel, the extended JLF also selects the candidate patches within a certain region in atlases.

**Sparse patch-base method**

Sparse learning methods are widely used in the field of image processing, e.g. face recognition. To take advantages of the sparse learning method, sparse patch-based method (SPBM) is proposed for multi-atlas segmentation. In SPBM, the target patch is reconstructed by the linear superposition of the selected candidate patches in the atlas images with sparsity constraint,

\[ \arg\min_{w_y} \frac{1}{2} ||P(y) - D(y)w_y||^2 + \lambda ||w_y||_1 \]  

(8)

where \( w_y = \{w(y, x_i) \}_{i=1, \ldots, N} \). \( M_y \) is the reconstruction coefficient vector, and \( D(y) = [P(x_{ik}) | x_{ik} \in N(y)] \) is a region-specific dictionary. The \( ||*||_1 \) is a \( l_1 \)-norm sparsity constraint, where \( \lambda \) controls the sparsity of the reconstruction coefficient vector \( w_y \). Due to the \( l_1 \)-norm constraint, only a small number of candidate patches in the atlas images with high similarity to the target patch will be selected for the subsequent label fusion process. Specifically, most of learnt reconstruction coefficients are zero. Then, the reconstruction coefficients are used as the voting weights to calculate the soft label of to-be-segmented voxel by Eq. (7) and get the final label by Eq. (3).

**Other state-of-the-art methods**

In the last decade, lots of label fusion methods are proposed. For example, several spare dictionary-based methods are proposed for label fusion. The sparse discriminative dictionary learning segmentation (DDLS) method is proposed for segmenting MR images. In DDLS, the dictionaries and classifiers are learnt simultaneously based the atlases, and

\[ \arg\min_{\alpha, \theta} \frac{1}{2} \|P - D\alpha\|_F^2 + \beta \|H - W\alpha\|_F^2 \]  

s.t. \( \|\alpha\|_1 \leq T \)  

(9)

**Figure 1.** Illustration of non-local strategies. The green and blue squares are target patch and the candidate patch, respectively. The blue boxes are the search region in the atlas images.
where \( P \) extract patches in the atlases image, \( D \) is the learned dictionary, and \( \alpha \) sparse coding coefficient matrix. \( H \) is the labels of the central voxels of the patches in \( P \), and \( W \) represents linear classifier. Finally, using the learnt dictionary and classifier for predicting the label of target voxel,

\[
l_y = \hat{W} \alpha,
\]

(10)

where \( l_y \) is the label vector of the target voxel. Also, the MAP criterion is used to obtain the final label of \( y \).

A multi-scale feature representation and label-specific patch partition hierarchical model is proposed for label fusion. In this method, the patches are represented with multiple scale, also the patches are divided into multiple pattern based on the label map of patches. As shown in Figure 2, the vectorized patches are divided into two group based on the label, the new vectors use 0 to replace the intensity values, if the voxel not belong to the label. Then, a sparse dictionary model with \(-\)norm and \(-\)norm constraints are used for learning the voting weight. Moreover, an iterative hierarchical method is used for label fusion, which gradually reduces the size of the patches to learn the voting weight by using the patches with non-zero voting weights in the last learning process.

A progressive dynamic dictionary method is proposed for label fusion in a layer-by-layer manner. Considering the gap between the image domain and the label domain, the authors use multi-layer dynamic dictionary to learn the representation of patches from image domain to label domain. As shown in Figure 3, this method uses a set of intermediate dictionaries to represent the image patches. By a sequence of intermediate dictionaries, the gap between the image domain and the label domain is reduced gradually. Finally, the method used the latest learnt intermediate dictionary to learn the voting weights for determining the final label of the to-be-segmented images. Notably, this method can be extended by most of multi-atlas methods.

A deep dictionary learning method is proposed for multi-atlas. The authors convert the conventional dictionary into a deep tree-like structure. Figure 4 shows a deep tree-like dictionary that consists of the candidate patch in the atlases. Specifically, the patches from the atlas are first divided into several groups based on the label of the center in the image patch. Then, a hierarchical k-means method is applied for each group to construct the remaining tree-like structure. Finally, the constructed deep tree-like dictionary is used for learning voting weight to determine the label of to-be-segmented by a hierarchical sparse patch label fusion.

In recent years, the deep learning methods achieve the success in many fields. The main advantage of the deep learning method is that it learns the feature representation from data without any pre-defined rules. Hence, several methods introduce the deep learning technique for learning the feature representation of image patch, and then use these learnt features for calculating the voting weight and label fusion. For instance, a high-order Boltzmann machine is introduced for learning the high-order representation of image patch. As shown in Figure 5, the high-order Boltzmann machine consists of two hidden units, i.e., the mean hidden units and covariance hidden units. Mean hidden units represent the mean intensity information of the image patch, while the covariance hidden units denote pair-wise dependencies information among voxels in the patch. In the training stage, the mean-co-

![Figure 2. The illustration of the label-specific atlas patch partition.](image)

![Figure 3. The framework of progressive dynamic dictionary based label fusion.](image)
The variance restricted Boltzmann machine (mcRBM) is trained based on the sampled image patches in the target images and atlas images. Then, in the segmentation stage, the patches are first represented by the mean feature and the covariance feature, then using these learnt high-order features for label fusion.

Also, the deep neural network is used for learning the representation of patches. As shown in Figure 6, this method extracts patches around the boundary of ROI, then uses the sampled image patches to train the neural network. Meanwhile, to learn the discriminative feature of the image patches, the deep neural network model also includes a discriminative term for making the learnt features discriminative for label fusion. In the test stage, the target patch in the target image and candidate patches in the atlas images are firstly converted to the learnt feature space by the learnt network, then the learnt features are used to calculate the voting weights for label fusion.

**DATASETS AND EVALUATION MATRICS**

There are several datasets that have been widely used in the study of multi-atlas based segmentation for brain ROI segmentation. We give a brief introduction of these dataset and also illustrate the evaluation metrics for assessing the segmentation performance of brain ROIs. Finally, we summarize the evaluation results of the state-of-the-art methods on the ADNI dataset.

**Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset**

The ADNI dataset includes 1.5T structural MRI and 3T structural MRI for hippocampus segmentation. There are three types of subject in the ADNI dataset, i.e., the Alzheimer’s disease (AD), the mild cognitive impairment (MCI) and the normal control (NC).

---

d. http://adni.loni.usc.edu/
Non-rigid Image Registration Evaluation Project (NIREP) dataset

NIREP dataset\(^7\) consists of 16 T1-weighted brain MR images, acquired from 8 normal male adults and 8 female adults. These MR images were obtained in a General Electric Signa scanner operating at 1.5 Tesla, using the following protocol: SPGR/50, TR 24ms, TE 7ms, NEX 1, matrix 256x192, FOV 24 cm. In this dataset, each MR image has been manually segmented into 32 ROIs.

LONI Probabilistic Brain Atlas (LPBA40) dataset

The LPBA40 dataset\(^8\) is provided by the Laboratory of Neuro Imaging (LONI) at UCLA. This dataset contains 40 brain images, and each image has 54 manually labeled ROIs. Specifically, these labeled ROIs were created manually to annotate brain structures of images in the LONI-LPBA40 dataset, and all MRI volumes were rigidly aligned to the MNI305 template.

IXI

The IXI dataset provides three type adult brain atlases, i.e., 30 subjects with 95 regions, 20 subjects with 83 regions and 20 subjects with 67 regions.

There are two common evaluation metrics used in assessing performance of segmentation. Specifically, the Dice ratio \((DR)\) measures the performance of segmentation as follows:

\[
DR = \frac{2|R_1 \cap R_2|}{|R_1| + |R_2|}
\]

where \(\cap\) denotes the overlap between automatically segmented regions \(R_1\) and \(R_2\), and \(|\cdot|\) denotes the number of voxels belonging to each ROI. The \(DR\) measures the degree of overlap between the segmented label map of image and the ground truth. The high value of \(DR\) can provide more precise quantitative of brain ROI volume.

Average surface distance \((ASD)\) is also used to measure the performance of segmentation algorithms.

\[
ASD = \frac{1}{2n_1n_2} \sum_{\tau_1 \in S(R_1)} d(\tau_1, S(R_2)) + \frac{1}{n_1n_2} \sum_{\tau_2 \in S(R_2)} d(\tau_2, S(R_1))
\]

where \(d(\cdot, \cdot)\) measures the Euclidean distance, and \(n_1\) and \(n_2\) are the numbers of vertices in surface \(S(R_1)\) and surface \(S(R_2)\), respectively. The \(ASD\) reflects the morphology similarity of between the segmented label map of image and the ground truth.

Based on the same dataset, performance of methods for segmenting brain MR images can be accessed with the above metrics. Table 1 shows the performance assessment results of the Dice ratio and average surface distance of the state-of-the-art methods using ADNI dataset for hippocampus segmentation. JLF\(^25\) considers the relationship between difference atlas, and reduces the possible risk of similar images. Hence, compared with LWV, JLF achieves better performance for hippocampus segmentation. Also, for the DDL\(^28\), leveraging the discriminative terms, it achieved better results than SPBM. The hierarchical model with multi-scale feature and label-specific patch partition\(^22\) reflects that using the multi-scale features and label information is helpful for multi-atlas segmentation. Progressive multi-atlas label fusion method\(^35\) shows that reduce the gap between image domain and label domain will increase the power of multi-atlas segmentation. The deep dictionary method\(^36\) uses deep tree-like dictionary to hierarchically organize the patch, and improve the representation power of sparse representation. The deep learning method\(^38\) also shows that using the learnt deep features will improve the segmentation performance of multi-atlas segmentation methods.

**APPLICATIONS AND PERSPECTIVES**

Segmentation is an important prerequisite step for medical image analysis. The multi-atlas based brain MR image segmentation aim to segment the whole brain into multiple ROIs and provide quantitative analysis of brain MR image. Makropoulos et al.\(^50\) segmented neonatal brain MR images into 50 ROIs based on multiple atlases, which provided a quantitative analysis of anatomical structure for subsequent research.

**Table 1. The performance assessment results of the state-of-the-art methods on ADNI dataset for hippocampus segmentation**

| Methods     | Subjects (n) | DR (%)   | ASD  | References No |
|-------------|--------------|----------|------|----------------|
| LWJoint     | 139          | 89.7(88.8) | -    | 25             |
| DDL         | 202          | 87.2     | -    | 28             |
| HLF         | 66           | 88.5     | 0.334| 32             |
| Progressive | 64           | 88.3     | -    | 35             |
| DSPBL       | 66           | 88.3     | 0.40 | 36             |
| NL1         | 135          | 86.45    | -    | 38             |

ANDI, Alzheimer’s disease neuroimaging initiative; DR, dice ratio; ASD, average surface distance.

e. http://www.nirep.org/
f. www.loni.usc.edu/atlases/  
g. http://brain-development.org/brain-atlases/adult-brain-atlases/
on the development of neonatal brain. Aljabar et al.\(^{51}\) segmented brain MR images via multi-atlas based segmentation method, then extracted the morphology features of brain anatomical structures based on the segmentation results, which were used in diagnosis of dementia. Additionally, the multi-atlas segmentation method is often applied in single ROI segmentation. For example, the hippocampus is related to the Alzheimer’s Disease (AD). Hence, to explore the relationship between the hippocampus and the AD, Pipitone et al.\(^{52}\) proposed a multi-atlas segmentation method for hippocampus segmentation by using few atlases, and provide a quantitative measurement of hippocampus.

During the last decade, a great amount of multi-atlas based methods have been proposed for brain ROI segmentation. The previous works have demonstrated that using the anatomical prior knowledge could help to segment the brain images which contain complicated structure, low intensity contrast around boundary of ROIs, large inter-subject variance and large inner-subject variance. However, the main disadvantage of multi-atlas based methods is the high computational cost. The voxel-by-voxel segmentation is very time-consuming for the high-dimension brain MR image, especially in the dictionary-based methods. The dictionary-based methods often construct the region-specific dictionary for each to-be-segmented voxels, which is very time-consuming.

In recent years, the deep learning methods have achieved great success in many fields. Although the deep learning methods are used in brain ROI segmentation, there are more researches that using the deep neural networks for learning the feature representation of the patches,\(^{37, 38}\) or use the patch-based network\(^{33}\) to predict the label map of patches directly. Hence, these methods also encounter the time-consuming issue. Meanwhile, due to the complicated structure of brain and the low intensity contrast around the boundary of ROI in brain MR images, the patch-based networks do not achieve a competitive result when compared with the multi-atlas methods. In the future, incorporating the anatomical prior into the end-to-end deep learning architectures, i.e., fully convolutional networks,\(^{54, 55}\) for brain ROI segmentation is an important direction.

**Conflict of interest disclosure**

All authors declared no conflict of interests.

**REFERENCES**

1. Devanand DP, Pradhaban G, Liu X, et al. Hippocampal and entorhinal atrophy in mild cognitive impairment - Prediction of Alzheimer disease. Neurology 2007; 68(11):828-36. doi: 10.1212/01.wnl.0000256697.20968.d7.

2. Zhang DQ, Wang YP, Zhou LP, et al. Multimodal classification of Alzheimer’s disease and mild cognitive impairment. Neuroimage 2011; 55(3):856-67. doi: 10.1016/j.neuroimage.2011.01.008.

3. Liu MX, Zhang DQ, Shen DG. Relationship induced multi-template learning for diagnosis of Alzheimer’s disease and mild cognitive impairment. IEEE Trans Med Imaging 2016; 35(6):1463-74. doi: 10.1109/Tmi.2016.2515021.

4. Jie B, Liu MX, Zhang DQ, et al. Sub-network kernels for measuring similarity of brain connectivity networks in disease diagnosis. IEEE Trans Image Process 2018; 27(5):2340-53. doi: 10.1109/Tip.2018.2799706.

5. Liu MX, Zhang J, Adeli E, et al. Landmark-based deep multi-instance learning for brain disease diagnosis. Med Image Anal 2018; 43:157-68. doi: 10.1016/j.media.2017.10.005.

6. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. Med Image Anal 2001; 5(2):143-56. doi: 10.1016/S1361-8415(98)00022-4.

7. Langerak TR, van der Heide UA, Kotte AN, et al. Improved label fusion in multi-atlas based segmentation by locally combining atlas selection and performance estimation. Comput Vision Image Understand 2015; 130:71-9. doi: 10.1016/j.cviu.2014.09.004.

8. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 2004; 23: S208-S19. doi: 10.1016/j.neuroimage.2004.07.051.

9. Avants BB, Epstein CL, Grossman M, et al. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 2008; 12(1):26-41. doi: 10.1016/j.media.2007.06.004.

10. Vercauteren T, Pennec X, Perchant A, et al. Diffomorphic demons: Efficient non-parametric image registration. Neuroimage 2009; 45(1):S61-S72. doi: 10.1016/j.neuroimage.2008.10.040.

11. Thirion JP. Image matching as a diffusion process: an analogy with Maxwell’s demons. Med Image Anal 1998; 2(3):243-60. doi: 10.1016/S1361-8415(98)00022-4.

12. Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. IEEE Trans Med Imaging 2004; 23(7):903-21. doi:
13. Heckemann RA, Hajnal JV, Aljabar P, et al. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. Neuroimage 2006; 33(1):115-26. doi: 10.1016/j.neuroimage.2006.05.061.

14. Artaechevarria X, Munoz-Barrutia A, Ortiz-de-Solorzano C. Combination strategies in multi-atlas image segmentation: application to brain MR data. IEEE Trans Med Imaging 2009; 28(8):1266-77. doi: 10.1109/Tmi.2009.2014372.

15. Aljabar P, Heckemann RA, Hammers A, et al. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. Neuroimage 2009; 46(3):726-38. doi: 10.1016/j.neuroimage.2009.02.018.

16. Lotjonen JMP, Wolz R, Koikkalainen JR, et al. Fast and robust multi-atlas segmentation of brain magnetic resonance images. Neuroimage 2010; 49(3):2352-65. doi: 10.1016/j.neuroimage.2009.10.026.

17. Sabuncu MR, Yeo BTT, Van Leemput K, et al. A Generative model for image segmentation based on label fusion. IEEE Trans Med Imaging 2010; 29(10):1714-29. doi: 10.1109/Tmi.2010.2050897.

18. Langerak TR, van der Heide UA, Kotte ANTJ, et al. Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE). IEEE Trans Med Imaging 2010; 29(12):2000-8. doi: 10.1109/Tmi.2010.2057442.

19. Lötjönen JM, Wolz R, Koikkalainen JR, et al. Fast and robust multi-atlas segmentation of brain magnetic resonance images. Neuroimage 2010; 49(3):2352-65.

20. Langerak TR, van der Heide VA, Kotte ANTJ, et al. Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE). IEEE Trans Med Imaging 2010; 29(12):2000-8. doi: 10.1109/Tmi.2010.2057442.

21. Zhang DQ, Wu GR, Jia HJ, et al. Confidence-guided sequential label Fusion for multi-atlas based segmentation. Int Conference on Medical Image Computing and Computer-Assisted Intervention; 2011 September 18-22; Toronto, Canada. Berlin, Heidelberg: Springer; 2011. p643-50. doi: 10.1007/978-3-642-23626-6_79.

22. Coupé P, Manjon JV, Fonov V, et al. Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. Neuroimage 2011; 54(2):940-54. doi: 10.1016/j.neuroimage.2010.09.018.

23. Rousseau F, Habas PA, Studholme C. A supervised patch-based approach for human brain labeling. IEEE Trans Med Imaging 2011; 30(10):1852-62. doi: 10.1109/Tmi.2011.2156806.

24. Zhang D, Guo Q, Wu G, et al., editors. Sparse patch-based label fusion for multi-atlas segmentation. International Workshop on Multimodal Brain Image Analysis. 2012 Oct 1-5; Nice, France. Berlin, Heidelberg: Springer; 2012. p94-102. doi: 10.1007/978-3-642-33530-3_8.

25. Wang HZ, Suh JW, Das SR, et al. Multi-atlas segmentation with joint label fusion. IEEE Trans Pattern Anal Machine Intell 2013; 35(3):611-23. doi: 10.1109/Tpami.2012.143.

26. Wang HZ, Yushkevich PA. Groupwise segmentation with multi-atlas joint label fusion. International Conference on Medical Image Computing and Computer-Assisted Intervention. 2013 Sep 22-26; Nagoya, Japan. Berlin, Heidelberg: Springer; 2013. p711-8. doi: 10.1007/978-3-642-40811-3_89.

27. Wang HZ, Yushkevich PA. Multi-atlas segmentation with joint label fusion and corrective learning — an open source implementation. Front Neuroinform 2013; 7. doi: 10.3389/fninf.2013.00027.

28. Tong T, Wolz R, Coupe P, et al. Segmentation of MR images via discriminative dictionary learning and sparse coding: Application to hippocampus labeling. Neuroimage 2013; 76(1):11-23. doi: 10.1016/j.neuroimage.2013.02.069.

29. Cardoso MJ, Leung K, Modat M, et al. STEPS: Similarity and Truth Estimation for Propagated Segmentations and its application to hippocampal segmentation and brain parcellation. Med Image Anal 2013; 17(6):671-84. doi: 10.1016/j.media.2013.02.006.

30. Wu GR, Wang Q, Zhang DQ, et al. A generative probability model of joint label fusion for multi-atlas based brain segmentation. Med Image Anal 2014; 18(6):881-90. doi: 10.1016/j.media.2013.10.013.

31. Sun L, Zu C, Zhang D. Reliability guided forward and backward patch-based method for multi-atlas segmentation. International Workshop on Patch-based Techniques in Medical Imaging; 2015 Oct 9; Munich, Germany. Cham: Springer; 2015. p128-36. doi: 10.1007/978-3-319-28194-0_16.

32. Wu GR, Kim MJ, Sanroma G, et al. Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition. Neuroimage 2015; 106:34-46. doi: 10.1016/j.neuroimage.2014.11.025.

33. Song Y, Wu G, Sun Q, et al. Progressive Label Fusion Framework for Multi-atlas Segmentation by Dictionary Evolution. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2015 October 5-9; Munich, Germany. Cham: Springer; 2015: p190-7. doi: 10.1007/978-3-319-24574-4_23.
34. Benkarim OM, Piella G, Ballester MAG, et al. Discriminative confidence estimation for probabilistic multi-atlas label fusion. Med Image Anal 2017; 42:274-87. doi: 10.1016/j.media.2017.08.008.
35. Song YT, Wu GR, Bahrami K, et al. Progressive multi-atlas label fusion by dictionary evolution. Med Image Anal 2017; 36:162-71. doi: 10.1016/j.media.2016.11.005.
36. Zu C, Wang ZX, Zhang DQ, et al. Robust multi-atlas label propagation by deep sparse representation. Pattern Recognition 2017; 63:511-7. doi: 10.1016/j.patcog.2016.09.028.
37. Liang S, Wei S, Zhang D, editors. High-order boltzmann machine-based unsupervised feature learning for multi-atlas segmentation. IEEE International Symposium on Biomedical Imaging. 2017 Apr 18-21; Melbourne, Australia. New York: IEEE; 2017. p507-10. doi: 10.1109/ISBI.2017.7950571.
38. Sanroma G, Benkarim OM, Piella G, et al. Learning non-linear patch embeddings with neural networks for label fusion. Med Image Anal 2018; 44:143-55. doi: 10.1016/j.media.2017.11.013.
39. Huo J, Wu J, Cao JW, et al. Supervoxel based method for multi-atlas segmentation of brain MR images. Neuroimage 2018; 175:201-14. doi: 10.1016/j.neuroimage.2018.04.001.
40. Sun L, Zu C, Shao W, et al. Reliability-based robust multi-atlas label fusion for brain MRI segmentation. Artif Intell Med 2019; 96:12-24. doi: 10.1016/j.artmed.2017.11.013.
41. Kittler J. Combining classifiers: A theoretical framework. Pattern Anal Appl 1998; 1(1):18-27. doi: 10.1007/Bf01238023.
42. Wang H, Yushkevich P. Multi-atlas segmentation with joint label fusion and corrective learning—an open source implementation. Front Neuroinform 2013; 7:27. doi: 10.3389/fninf.2013.00027.
43. Zhang QA, Li BX. Discriminative K-SVD for Dictionary Learning in Face Recognition. IEEE Conference on Computer Vision and Pattern Recognition. 2010 Jun 13-18; San Francisco, CA, USA. New York: IEEE; 2010. p2691-8. doi: 10.1109/Cvpr.2010.5539989.
44. Yang M, Zhang L, Feng XC, et al. Fisher discrimination dictionary learning for sparse representation. International Conference on Computer Vision. 2011 Nov 6-13; Barcelona, Spain. New York: IEEE; 2011. p543-50. doi: 10.1109/ICCV.2011.6126286.
45. Ranzato M, Hinton GE. Modeling pixel means and covariances using factorized third-order boltzmann machines. IEEE Conference on Computer Vision and Pattern Recognition. 2010 Jun 13-18; San Francisco, CA, USA. New York: IEEE; 2010. p2551-8. doi: 10.1109/Cvpr.2010.5539962.
46. Jack CR, Bernstein MA, Fox NC, et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J Magn Resonan Imaging 2008; 27(4):685-91. doi: 10.1002/jmri.21049.
47. Christensen GE, Geng X, Kuhl JG, et al. Introduction to the non-rigid image registration evaluation project (NIREP). IEEE Trans Magnet 2006; 30(5):2972-5. doi: 10.1109/20.312561.
48. Shattuck DW, Mirza M, Adisetiyo V, et al. Construction of a 3D probabilistic atlas of human cortical structures. Neuroimage 2008; 39(3):1064-80. doi: 10.1016/j.neuroimage.2007.09.031.
49. Evans AC, Collins DL, Mills SR, et al. 3D Statistical Neuroanatomical Models from 305 MRI Volumes. IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference. 1993 Oct 31 - Nov 6; San Francisco, CA, USA. New York: IEEE; 1993. 1813-7. doi: 10.1109/Nssmic.1993.373602.
50. Makropoulos A, Gousias IS, Ledig C, et al. Automatic whole brain MRI segmentation of the developing neonatal brain. IEEE Trans Med Imaging 2014; 33(9): 1818-31. doi: 10.1109/Tmi.2014.2322280.
51. Aljabar P, Rueckert D, Crum WR. Automated morphological analysis of magnetic resonance brain imaging using spectral analysis. NeuroImage 2008; 43(2): 225-35. doi: 10.1016/j.neuroimage.2008.07.055.
52. Pipitone J, Park MTM, Winterburn J, et al. Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. Neuroimage 2014; 101:494-512. doi: 10.1016/j.neuroimage.2014.04.054.
53. Chen YJ, Gao HY, Cai L, et al. Voxel Deconvolutional Networks for 3D Brain Image Labeling. Kdd’18: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018 Aug 19-23; London, UK. New York: ACM; 2018. p1226-34. doi: 10.1145/3219819.3219974.
54. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention. 2015 Oct 5-9; Munich, Germany. Cham: Springer; 2015. p234-41. doi: 10.1007/978-3-319-24574-4_28.
55. Shelhamer E, Long J, Darrell T. Fully Convolutional Networks for Semantic Segmentation. IEEE Trans Pattern Anal Machine Intell 2017; 39(4):640-51. doi: 10.1109/TPami.2016.2572683.