A robust discriminative multi-atlas label fusion method for hippocampus segmentation from MR image

Wenna Wang [a,b,c], Xiuwei Zhang [a,b,c,*], Yu Ma [d], Hengfei Cui [a,b,c], Rui Xia [d,e], Yanning Zhang [a,b,c]

[a] School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, China
[b] Shaanxi Provincial Key Lab. of Speech and Image Information Processing (SAIIP), School of Computer Science, Northwestern Polytechnical University, Xi'an, China
[c] National Engineering Laboratory for Air-Sea-Earth-Sea Integrated Big Data Application Technology, China
[d] School of Ningxia University, Yinchuan 750021, China
[e] Zhejiang Dahua Technology Co., Ltd, Hangzhou 310000, China

ARTICLE INFO

ABSTRACT

Accurate and automatic segmentation of the hippocampus plays a vital role in the diagnosis and treatment of nervous system diseases. However, due to the anatomical variability of different subjects, the registered atlas images are not always perfectly aligned with the target image. This makes the segmentation of the hippocampus still face great challenges. In this paper, we propose a robust discriminative label fusion method under the multi-atlas framework. It is a patch embedding label fusion method based on conditional random field (CRF) model that integrates the metric learning and the graph cuts by an integrated formulation. Unlike most current label fusion methods with fixed (non-learning) distance metrics, a novel distance metric learning is presented to enhance discriminative observation and embed it into the unary potential function. In particular, Bayesian inference is utilized to extend a classic distance metric learning, in which large margin constraints are instead of pairwise constraints to obtain a more robust distance metric. And the pairwise homogeneity is fully considered in the spatial prior term based on classification labels and voxel intensity. The resulting integrated formulation is globally minimized by the efficient graph cuts algorithm. Further, sparse patch based method is utilized to polish the obtained segmentation results in label space. The proposed method is evaluated on IABA dataset and ADNI dataset for hippocampus segmentation. The Dice scores achieved by our method are 87.2%, 87.8%, 88.2% and 88.9% on left and right hippocampus on both two datasets, while the best Dice scores obtained by other methods are 86.0%, 86.9%, 86.8% and 88.0% on IABA dataset and ADNI dataset respectively. Experiments show that our approach achieves higher accuracy than state-of-the-art methods. We hope the proposed model can be transferred to combine with other promising distance measurement algorithms.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

As we all known, hippocampus plays a vital role in human brain. It has been found that the hippocampus is associated with a variety of brain neural diseases, including Alzheimer's disease (AD), geriatric depression, etc. [1]. Different diseases have various degrees of atrophy. For example, patients with AD have significant atrophy in the CA1 and ERC sub-regions of hippocampus, while in elderly patients with depression, the volume of SUB and CA2 sub-regions decrease notably. Early detection of changes in the hippocampus is essential for the early prevention and diagnosis of neural diseases. Hence, the automatic and accurate extraction of hippocampus from Magnetic Resonance (MR) images has become a pivotal task in medical image analysis [2,3].

Due to the irregular shape of the hippocampus and the blurred boundary with surrounding tissues, the automatic and accurate segmentation of the hippocampus is a challenging task. The earliest segmentation of specific anatomical structures is achieved by expert manual labeling. But there are long-term and error-prone defects, which limit the application of manual marking in big data [4–6]. Therefore, the automatic segmentation of MR images has become a hot topic in medical image analysis. Recently, multi-atlas segmentation (MAS) method receives broad attention as its good performance in medical image segmentation [7–12]. The MAS method, which integrates the prior knowledge of medical atlas into the segmentation process and then combines with the efficient label fusion algorithm, can obtain a quite accurate segmentation re-

---

* Corresponding author at: School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, China.

sult on brain MR images. Typically, the MAS method consists of two main steps: (1) atlas image registration, and (2) label fusion [13]. Specifically, in the first step, the target image and each candidate image are spatially converted into the identical space to build a spatial correspondence [14,15]. Then in the label fusion step, the labels from different atlas images are fused to obtain the final estimation of the target image [16–18].

In the study of MAS, many scholars work on labeling methods to improve the segmentation performance [19–29]. Since the anatomical variability of various subjects, there are different degrees of registration errors between atlas images and target image. However, appropriate label fusion strategies can effectively reduce the impact of the unavoidable registration error. The existing label fusion methods for medical image segmentation can be roughly divided into three categories: voting-based labeling methods, patch-based labeling methods, and deep learning-based labeling methods. Majority voting (MV) is one of the widely used voting-based label fusion methods [19]. It assigns the same weight for all atlas. There are also other techniques such as STAPLE, which estimates the weights by rater statistics [20,21]. However, these approaches are based on one assumption that there is a voxel-to-voxel correspondence between the atlas image and the target image. It makes the segmentation results sensitive to the registration errors. The non local patch-based labeling method (PBM) [22] can solve this problem. In PBM, voting weights are computed between the target image patch and atlas patches in a fixed search region. More recently, sparse patch-based method (SPBM) shows improved robustness and accuracy of segmentation results. In this approach, candidate atlas patches with sparse constraints are applied to obtain the weight of the target patch [6,23,24]. Some approaches are also proposed based on PBM/SPBM to improve the accuracy and robustness. For example, AR+LKSRC method constructs a probabilistic atlas of deep structures by registration and applies a sparse representative classifier to refine results around boundaries [25]. There is also another framework that can improve the discriminative power in patch-based label fusion by a manner of using neural networks to calculate patch embeddings [21]. An appealing aspect is that deep learning methods can automatically learn high level features of images through manual design or automatic search of neural networks. However, the memory complexity of training limits the number of atlases which can be effectively used for segmentation. For example, a correlation method uses only 5 atlas images for each target image by using convolutional neural networks [21,26]. By contrast, in brain MR image segmentation, usually around 10 atlases are used [27,28]. So we focus on patch-based label fusion method under multi-atlas framework. It is worth noting that a formula is integrated to embed the patch into the CRF model in our work.

In the literature, most existing PBM methods [19–28] adopt a predefined distance metric model to measure the intensity-based similarity between target patch and atlas patches. This kind of similarity measure can not effectively characterize statistical distributions of image intensities [29]. For example, patches with similar intensity values may be compatible with different segmentation labels [30,31]. To alleviate this drawback, a robust discriminative label fusion method under MAS framework is proposed. Instead of a pre-fixed distance metric model, a novel metric learning is integrated into CRF model with the graph cuts to enhance the discriminative observation. Furthermore, to fully make use of the prior information of the multi-atlas, SPBM method is adopted to polish the obtained segmentation in label space. The proposed method achieves more accuracy segmentation results. The main contributions are as follows:

(1) A fast "coarse-fine" hybrid registration method is presented to alleviate the high computational cost of image registration.

The proposed method adopts the resampling algorithm in the coarse registration stage. Compared with rigid registration, resampling does not require any iterative optimization process to achieve rough alignment between each candidate image and the target image. Thereby the computational cost is reduced compared with classical "coarse-fine" registration stage.

(2) A robust discriminative multi-atlas label fusion method is proposed. It is a patch embedding segmentation method on the basis of CRF model, which integrates distance metric learning and graph cuts. Specifically, Bayesian theorem is adopted to extend the classical distance metric learning and large margin constraints are exploited instead of pairwise constraints. Compared with the fixed (non-learning) distance metrics, it can obtain a more robust distance metric. This well-designed learning metric is adopted to model the unary potential function of CRF. The spatial prior term fully considers the pairwise homogeneity of two adjacent nodes and is used in the pairwise potentials function of CRF. Both parts can enhance the discriminative observations, so this method potentially improves the segmentation performance of the hippocampus.

(3) Furthermore, considering the label map with rich prior information is not fully used, we refine the segmentation results based on SPBM method. In this approach, candidate label patches with sparse constraints are applied to obtain the weight of the initial segmentation patch. Compared with state-of-the-art methods, the proposed approach can obtain more accurate hippocampal segmentation results.

The rest of the paper is arranged as follows. The fast "coarse-fine" hybrid registration method and the proposed robust discriminative multi-atlas label fusion method are presented in Section 2; Section 3 describes the experimental results for the hippocampus segmentation; The discussion and conclusions are given in Section 4 and Section 5 respectively.

2. Materials and methods

MAS aims to achieve the segmentation of the target subject by fully utilizing multiple sets of atlas grayscale images and their corresponding label maps. The proposed segmentation scheme is based on MAS and involves four main steps: (1) MR images pre-processing, (2) "coarse-fine" hybrid registration, (3) label fusion based on our robust discriminative multi-atlas label fusion method and (4) refine segmentation by SPBM in label space. Fig. 1 shows a flowchart of the proposed overall robust multi-atlas label fusion method.

When analyzing the structure of the brain, many factors can interfere with image segmentation, such as the background, the brain shell, and the gray unevenness of the image. Therefore, it is essential to pre-process the primitive MR images before segmenting the hippocampus. In our work, skull stripping, N4-based bias field correction and intensity standardization are performed, thereby standardizing the strength range of experimental data, so that the subsequent processing can be carried out smoothly. In our method, 8 most similar MR images are selected for each target subject from the data set based on the mutual information (MI) measurement. Then, each candidate image is registered with the target image by the proposed fast "coarse-fine" hybrid registration. Next, based on the registered atlases, the target label is inferred by the proposed robust discriminative multi-atlas label fusion (RDLF) method. Furthermore, SPBM method is adopted to refine the segmentation results in label space (RRDLF), so as to fully use the prior information of atlases. The specific details will be introduced in the following sections.
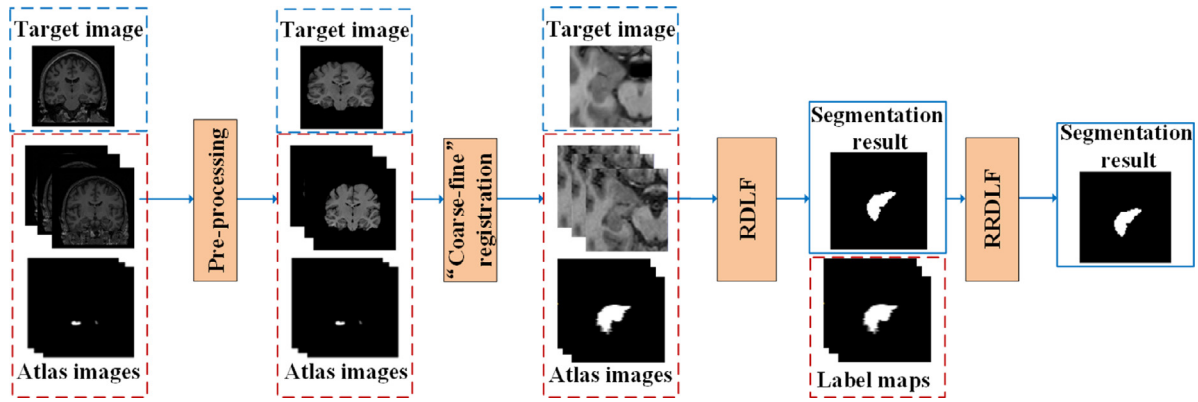
Fig. 1. Flow chart of the proposed hippocampus segmentation scheme. The proposed segmentation scheme involves four main steps: (1) MR images preprocessing, (2) "coarse-fine" hybrid registration, (3) label fusion based on our robust discriminative multi-atlas label fusion method and (4) refine segmentation by SPBM in label space.
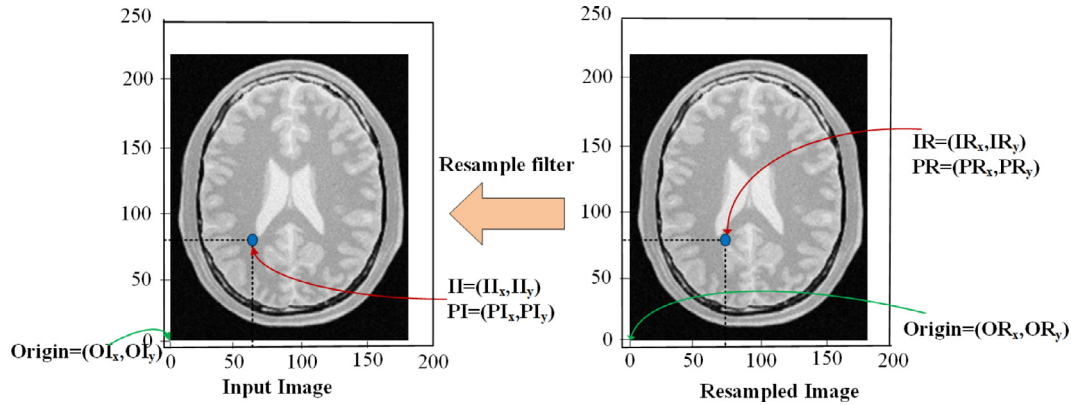

Fig. 2. Schematic of resampling. After resampling, the origin, spacing, and size of the input image are changed to become the same as the reference image.

2.1. A fast "Coarse-fine" hybrid registration

"Coarse-fine" hybrid registration is utilized in most multi-atlas segmentation algorithms to establish spatial correspondence between the atlas and the target image. The fine registration, generally referring to non-rigid registration, requires that the atlases and the target image possess the identical spatial sampling point and the same spatial distance of the sampling points. Therefore, a coarse registration is necessary before the fine registration process to reduce the deformation difference of the non-rigid registration. Rigid registration is the most commonly used coarse processing method. However, rigid registration is time-consuming, since the data amount of each 3D medical image is very large and rigid registration needs to iteratively calculate transformation parameters between dozens of atlas images and target image. Hence, it makes it difficult to meet clinical needs.

In this article, we propose a fast "coarse-fine" hybrid registration using resampling instead of rigid registration. The resampling method maps the spatial coordinates of the input image to generate a new image based on the voxel position and voxel spacing of the reference image. Thus the spatial sampling point and sampling distance of the input image are adjusted. As shown in Fig. 2, it is a simple schematic of the resampling process with $2D$ slices as an example to make it easier to understand. The coordinate of pixel $IR = (IR_x, IR_y)$ of the resampled image in physical space is $PR = (PR_x, PR_y)$. Since resampling is performed in spatial coordinates, not pixel/grid coordinates. Therefore, the resampling filter adjusts the gray of the output image according to the brightness value of the spatial position PI of the input image. The coordinate of the PI point is associated with pixel $II = (II_x, II_y)$ on the input image. If II does not fall on the grid position, the value pointing

to the output pixel will be calculated by inserting values around the non-integer IR in the input image. By traversing each point in the reference image, the input image is resampled to the template space. Thus the problem of inconsistency of space sampling points between atlas image and the target image is adjusted.

Unlike rigid registration, the resampling operation makes the atlas images have the same isotropic sampling rate as the reference image, while there is no complicated iterative optimization process. Meanwhile, the size and center of each candidate image is regulated to be consistent with the target image. Thus, the roughly alignment can be achieved in a faster way. In order to increase the speed of MAS and reduce the impact of other anatomical structures on hippocampus segmentation, here we define the region of interest for the anatomic structure as the smallest bounding box, which contains the research structure (left or right hippocampus) of all training atlases. After that, the standardized object images are cut around the structures of interest. For the fine registration task, the differential homeomorphic Demons algorithm [32] is adopted on account of its high efficiency, high accuracy and the ability of dealing with large deformation problems. The overall framework of our fast "coarse-fine" hybrid registration is shown in Fig. 3.

2.2. Robust discriminative multi-atlas label fusion method

In this subsection, the robust discriminative label fusion method is elaborated in detail. Fig. 4 depicts the main idea of our label fusion method. In the proposed multi-atlas framework, we combine a novel metric learning and graph cuts algorithm by an effective segmentation scheme based on CRF model. The presented distance metric learning is used to model the unary potential. It
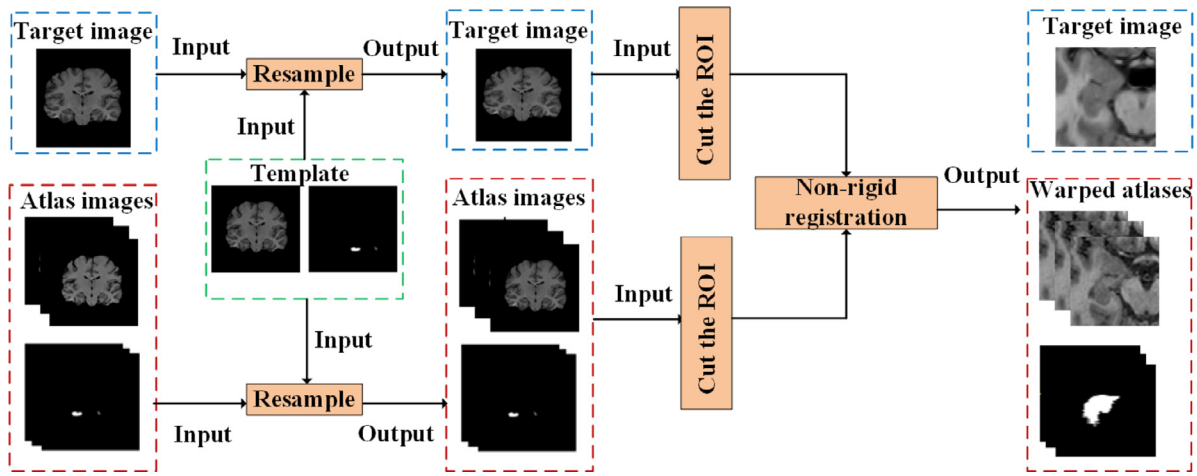
Fig. 3. The flowchart summarizes the process of the fast "coarse-fine" hybrid registration. In the coarse registration phase, the resampling algorithm is used to adjust the problem of inconsistent of sampling points, thus realizing the rough alignment of the atlas images and the target image. In the non-rigid registration phase, the differential homeometry Demons algorithm is applied to obtain accurate registration results.
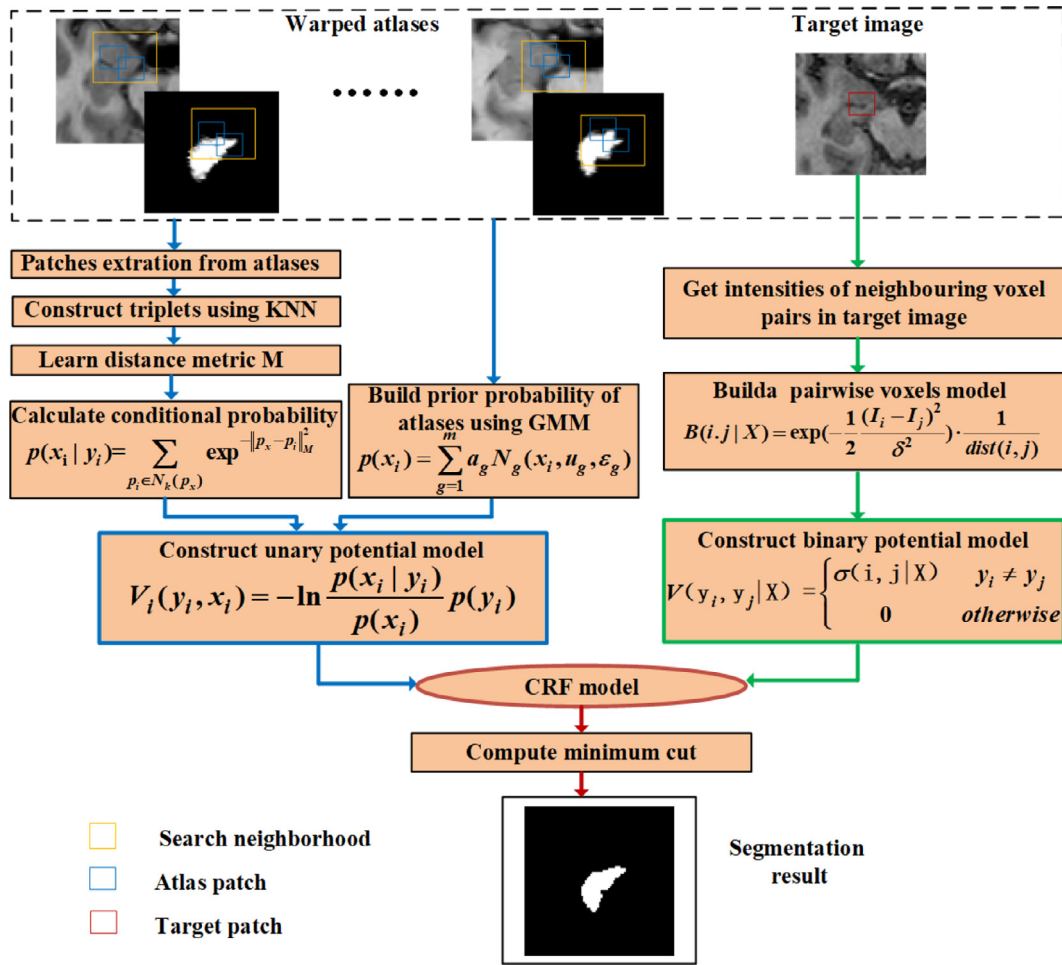

Fig. 4. Illustration of robust discriminative multi-atlas label fusion method for hippocampus segmentation. The presented distance metric learning method is used to model the unary potential function (left part). The graph cuts part is designed to model the pairwise potentials function of CRF model (right part).

employs a Bayesian inference to extend a classic distance metric learning and uses large margin constraints instead of the pairwise constraints. Spatial prior term is designed to model the pairwise potentials function of CRF model. This part considers the pairwise homogeneity in classification labels and the pixel intensity. The resulting segmentation problem is described as a mapping esti-

mation for searching the most likely binary image segmentation, which allows the solution to be calculated effectively through the maximum flow or minimum cut optimization process. The specific details of the whole model are given as bellows:

Assume that X is the observation vector of the entire image (i.e. the characteristics of all pixels), and $Y = \{y_1, y_2, \ldots, y_i \ldots\}$ is the

category label corresponding to all pixels in the image X. For two types of segmentation problems, the label of the pixel at index i is $y_i \in Y$ and $y_i \in \{0, 1\}$, where 0 represents the background and 1 represents the target.

In CRF model, the posterior probability is directly modeled as a Gibbs distribution according to the Hammersley-Clifford theorem. When the image observation data X is given, CRF directly models the posterior probability distribution of the label Y. The formula is as follows:

$$P(Y|X) = \frac{1}{N(X)} exp(-E(Y|X)), \qquad (1)$$

where N is the normalization factor, and $E(Y|X)$ is the Gibbs energy function. In general, the problem of segmenting an image can be viewed as the maximum posterior estimation problem of searching for the most likely label mapping:

$$Y^* = \arg \max_Y P(Y|X) = \arg \min_Y E(Y|X). \qquad (2)$$

Using CRF for image segmentation, assume only the unary potential and the pairwise potentials function are considered, the energy function can be expressed as follows:

$$E(Y|X) = \sum_{i \in \Omega} V_i(y_i, x_i) + \frac{1}{\zeta} \sum_{i \in N_i} \sum_{j \in N_i} V_{ij}(y_i, y_j, X), \qquad (3)$$

where Ω is the index set of the voxel x in the image X, and N_i is the index set of the neighborhood of the voxel x_i. $V_i(y_i, x_i)$ is the unary potential function, which is used to describe the probability that the voxel with index i belongs to the label y_i. $V_{ij}(y_i, y_j, X)$ is the pairwise potentials function that describes the association between adjacent points of a voxel and its impact on category determination. ζ measures the relative importance of the energy of the second-order potential function. In the cases that the feature vector values of two neighboring pixels are the same but different labels are assigned, the pairwise potentials function should impose a larger penalty.

2.2.1. The unary potential

This paper uses Bayesian formula to extend the unary potential function of the CRF model, and the probability value of voxel features x_i that belongs to the corresponding category is used as the value of the unary potential energy function:

$$V_i(y_i, x_i) = -lnp(y_i|x_i) = -ln\frac{p(x_i|y_i)p(y_i)}{p(x_i)}, \qquad (4)$$

where $x_i \in X$ is the feature vector. The gray level of each region of medical image changes relatively slow and the gray statistical histogram always shows multi-peak characteristics. Compared with single Gaussian model, the Gaussian mixture model (GMM) can reduce the misclassification of pixels. Hence, we adopt Gaussian mixture model to describe the gray statistical characteristics of MR images:

$$p(x_i) = \sum_{g=1}^{m} \alpha_g N_g(x_i, u_g, \varepsilon_g), \qquad (5)$$

where m represents the number of Gaussian model components, $\sum_{g=1}^{m} \alpha_g = 1$, α_g is a weighting factor, N_g is a Gaussian probability density function, which can be expressed as follows:

$$N_g(x_i, u_g, \varepsilon_g) = \frac{1}{\sqrt{(2\pi)^g|\varepsilon_g|}} \exp(-\frac{1}{2}(x_i - \mu_g)^T \varepsilon_g^{-1}(x_i - \mu_g)), \quad (6)$$

where u_g and ε_g represent the mean vector and covariance matrix of all pixels belonging to the gth model. In this paper, we adopt Expectation Maximum algorithm to estimate GMM parameters.

Then, the focus is on how to get a robust posterior probability estimation. The following formula is used to calculate the image likelihoods of each voxel in target image:

$$p(x_i|y_i) \propto \sum_{p_i \in N_k(p_x)} exp^{-\|p_x - p_i\|_M^2}, \qquad (7)$$

where p_x is the target patch with the size of $r_p \times r_p \times r_p$. $N_k(p_x)$ is the neighborhood which is similar to the target patch p_x. For each target voxel, voxels are extracted from per atlas image in a fixed cubic search region with size of $r_s \times r_s \times r_s$. Atlas patches which are centered at these voxels are used to construct a patch library $PL = (p_i, y_i)_{i=1,2,...,N}$, in which p_i is the ith image patch and y_i is the label of its center voxel. Note that M is the distance metric, and the distance between target patch p_x and atlas patch p_i is computed by Mahalanobis distance. It's calculation formula is as follows:

$$d_{xi}^2 = d(p_x, p_i) = (p_x - p_i)^T M(p_x - p_i). \qquad (8)$$

According to $d(p_x, p_i)$, k nearest neighbor algorithm is used to select k nearest training samples, and then form a nearest neighborhood set $N_k(p_x)$. So the unary potential energy function is defined as:

$$V_i(y_i, x_i) = -lnp(y_i|x_i) = -ln\frac{\sum_{p_i \in N_k(p_x)} exp^{-\|p_x - p_i\|_M^2}}{\sum_{g=1}^{m} \alpha_g N_g(x_i, u_g, \varepsilon_g)} p(y_i). \qquad (9)$$

In order to gain a more robust distance M, we develop a new metric learning method in this work. Following reference [33,34] to learn a robust distance metric, in this work, the Bayesian inference is applied to distance metric learning, which estimates the posterior distribution with large margin nearest neighbor constraints. According to Bayesian inference, prior probability of the distance matrix parameter M is given as $p(M)$, and likelihood probability is defined as $p(A|M)$. Then the posterior distribution $p(M|A)$ can be estimated based on the training data A. For our work, we introduce a Gaussian prior for the transformation matrix M. According to the principle of large margin nearest neighbor method, the likelihood function is defined as:

$$P(A|M) = \prod_{i,j,l \in A} prob(p_i, p_j, p_l, y_i, y_j, y_l|M)$$
$$= C \prod_{i,j,l \in A} exp[-2 \times max(1 + d_{ij}^2 - d_{il}^2, 0)]. \qquad (10)$$

Let $y_{ij} = 1$ denotes a pair of similar data samples with $y_i = y_j$, and $y_{ij} = 0$ indicates a dissimilar pair. In the above formula, A is the training set containing $|A|$ independent triplets (i, j, l), which satisfies $1 + d_{ij}^2 - d_{il}^2 > 0$, $(y_{ij})(1 - y_{il}) = 1$, p_j and p_l belong to the set of neighbors of p_i. The triplets are constructed by KNN method, in which we select the nearest homogeneous sample and the nearest heterogeneous sample for each sample from the patch library PL. Here C is a normalizing constant. The function incorporates the idea of a margin, in which the purpose is to learn a distance metric that can separate d_{ij} and d_{il} by a large margin.

According to the formula in the Bayesian model, the posterior distribution probability $P(M|A)$ of the transformation matrix parameter M is satisfied $P(M|A) \propto P(A|M)P(M)$. In order to use Bayesian theorem, a lot of Gaussian distributions are worked to approximate each single likelihood function. The Gaussian approximation of the known Laplace distribution is described as follows:

$$La(x|0, \delta) = \frac{1}{2\delta} exp- \frac{|x|}{\delta}$$
$$= \int_0^\infty \frac{1}{\sqrt{2\pi\beta}} exp(-\frac{x^2}{2\beta}) \cdot \frac{1}{2}\delta^{-2} exp(-\frac{\beta}{2\delta^2}) d\beta. \qquad (11)$$

Since:

$$exp(-2 \times max(\alpha, 0)) = exp(-|\alpha| - \alpha)$$
$$= \int_0^\infty \frac{1}{\sqrt{2\pi\beta}} exp(-\frac{\alpha^2}{2\beta} - \alpha)$$
$$\cdot \frac{1}{2} exp(-\frac{\beta}{2}) d\beta. \qquad (12)$$

So the likelihood function becomes:

$$P(A|M) = \prod_{i,j,l \in A} prob(p_i, p_j, p_l, y_i, y_j, y_l | M)$$
$$= \int_0^\infty \frac{1}{\sqrt{2\pi\lambda_{ijl}}} exp(-\frac{1}{2} \frac{(1 + d_{ij}^2 - d_{il}^2 + \lambda_{ijl})^2}{\lambda_{ijl}}) d\lambda_{ijl}, \quad (13)$$

where λ_{ijl} is essentially an induced parameter. M is the distance matrix. In our work, the distance function d_{ij}^2 are reformulated in a linear form, namely:

$$d_{ij}^2 = Tr[M(p_i - p_j)(p_i - p_j)^T] = \gamma^T p_{ij}, \qquad (14)$$

where the γ and p_{ij} are the vectorized forms of matrices M and $(p_i - p_j)(p_i - p_j)^T$ respectively. So the objective can be boiled down to seeking the optimal γ, whose posterior distribution can be estimated by employing the factorial variation inference [35–37]. The main idea is to adopt factorized variational distribution to approximate the distribution of the posterior probability and then the goal becomes to minimize the KL divergence until convergence. The parameter λ_{ijl} is initialized in the same way as factorized variational distribution. More detailed information on this part can be found in reference [37,38]. With the learned Mahalanobis distance metric M, the unary potential function in our work can be easily constructed.

2.2.2. The pairwise potentials

The pairwise potentials function consists of a spatial prior term, which considers the pairwise homogeneity based on multi-atlas voxel intensities and classification labels. Following Boykov's analysis in the article [39], the function is defined as follows:

$$V_{ij}(y_i, y_j | \mathbf{X}) = \begin{cases} B(i, j|\mathbf{X}) & y_i \neq y_j \\ 0 & otherwise, \end{cases} \qquad (15)$$

$$B(i, j|\mathbf{X}) = exp(-\frac{1}{2} \frac{(I_i - I_j)^2}{\sigma^2}) \cdot \frac{1}{dist(i, j)}, \qquad (16)$$

where $B(i, j|\mathbf{X})$ represents the cost of target boundaries. I_i and I_j are the intensity values of voxels with index i and j. If the voxels jumps sharply between I_i and I_j, they are considered to belong to different classes and only small penalties are assigned. $dist(i, j)$ is the Euclidean distance. The parameter δ represents the noise estimation. The pairwise potentials function emphasizes homogeneous segmentation between adjacent voxels. It prevents the over segmentation of sharp edges by weighting the penalty for heterogeneity in terms of intensity similarities of the related voxels.

With the equality

$$Y^* = \arg\max_Y P(Y|X) = \arg\min_Y E(Y|X)$$
$$= \arg\min_Y \sum_{i \in \Omega} (-lnp(y_i|x_i)) + \frac{1}{\zeta} \sum_{i \in N_i} \sum_{j \in N_i} V_{ij}(y_i, y_j, X), \qquad (17)$$

the original optimization problem can be converted into an energy minimization problem, which can be optimized in a proper form by graph cutting algorithm [39,40].

2.3. Refine segmentation result in label space

In Section 2.2, the atlases are utilized to calculate Gaussian mixture model and estimate the posterior probability. Actually, the label information of atlases is rich of prior information and is very important for similarity measurement. To take fully advantage of these prior knowledge, on the basis of the segmentation result obtained in the previous subsection, we adopt SPBM to further refine the segmentation in label space. The principle of SPBM is shown in Fig. 5. In this subsection, we will present the processing of the segmentation refinement.

The PBM model can be depicted by Eq. (2.3). Suppose v_i represents the voxel at location i of the image and $v_{s,j}$ represents the voxel of the subject s at location j. The final label is estimated based on the weighted label fusion of each selected sample. The formula is as follows

$$voting(v_i) = \frac{\sum_{s=1}^N \sum_{j \in O_i} w_{s,j}^i y_{s,j}}{\sum_{s=1}^N \sum_{j \in O_i} w_{s,j}^i}. \qquad (18)$$

where O_i is the search area for N selected atlases, $y_{s,j}$ is the label of voxel $v_{s,j}$. $w_{s,j}^i$ is the corresponding weight between v_i and $v_{s,j}$, which is often calculated by the similarity between two voxels. Obviously, the similarity measure in Eq. (2.3) is the core of the PBM. Inspired by sparse representation in face recognition and other aspects, adding a sparse constraint to a fusion algorithm based on non-local averages can get more accurate segmentation results [31]. Therefore, the following sparse representation model is adopted to calculate the weight $w_{s,j}^i$ between two different labels y_i and $y_{s,j}$.

$$\min_{w_{s,j}^i} \frac{1}{2} \left\| \sum_{s=1}^N \sum_{j \in O_i} D_s w_{s,j}^i - G(y) \right\|_2^2 + \lambda \sum_{s=1}^N \sum_{j \in O_i} \left\| w_{s,j}^i \right\|_1. \qquad (19)$$

Eq. (18) is similar to the traditional SPBM algorithm, except that the elements of D_s and $G(y)$ are label information instead of intensity information. The regularization parameter λ controls the sparsity of $w_{s,j}^i$. Specifically, the previous item is used to minimize reconstruction errors. The function of the latter term is a sparse constraint, where λ is a regularization parameter balancing the relative contributions of these two terms. Eq. (18) is a $l_1 - norm$ optimisation problem, so the sparse representations can be obtained by using the Lasso method [41]. Once the weighting vectors are calculated, the estimation of the label can be derived using Eq. (2.3).

3. Experiments

The performance of the proposed methods are evaluated on two brain MR image datasets for right and left hippocampus segmentation. The experimental results will be introduced in details in the following paragraphs.

3.1. Dataset

The proposed methods are evaluated on Individual Adult Brain Atlases (IABA) dataset and Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, which are widely used for evaluating hippocampus segmentation. IABA dataset[1] is launched by the Medical Brain Research Dataset at Imperial College London. It contains 20 samples, each of which is provided as a T1-weighted MR grayscale image and a corresponding label map. Experts labeled 67 structures in labeled brain images, such as the hippocampus, thalamus, cerebellum, and tonsils. Figure 6 is the slices display of the MR

---

[1] http://brain-development.org/brain-atlases/adult-brain-atlases/individual-adult-brain-atlases/
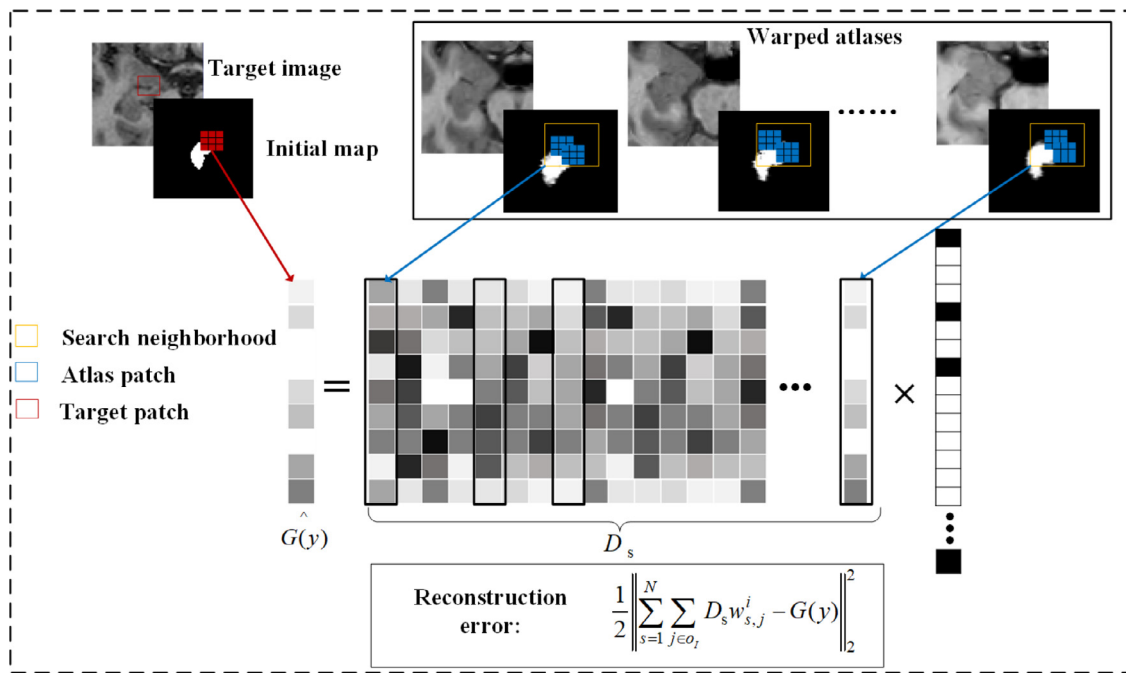
Fig. 5. Overview of SPBM which embeds the label structure (take a 2D image as an example, and assume the patch size is 3×3).

The reconstruction error shown in the figure:

$$\text{Reconstruction error:} \quad \frac{1}{2}\left\|\sum_{s=1}^{N}\sum_{j\in o_l} D_s w_{s,j}^i - G(y)\right\|_2^2$$
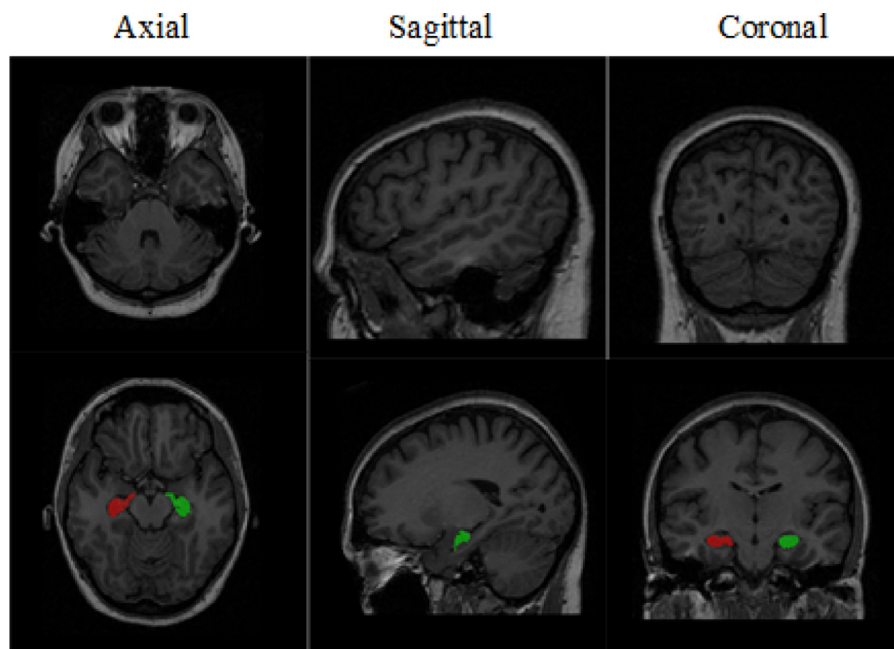

Fig. 6. The axial, sagittal, and coronal slice display of the MR image. The first row is the two-dimensional display of the 43rd slice, and the second row is the two-dimensional display results of the 56th, 70th, 105th slices, respectively. The red part represents the right hippocampus, and the green indicates the left hippocampus. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

image available from IBAB dataset. ADNI dataset[2] is a large clinical medical image dataset, which enables all data and samples to be shared with scientists around the world [42–44]. It contains a lot of annotated MR images, from which we randomly select 105 sets of MR images. Demographic and clinical information of these subjects are summarized in Table 1.

Table 1
Demographic and clinical information of the studied ADNI subjects.

	Age	Male/Female	MMSE	Education
NC	75.8±6.8	17/13	29.1±1.1	15.2±3.1
MCI	75.2±7.6	22/15	26.3±2.8	15.7±2.6
AD	74.0±8.1	20/18	21.9±4.1	15.9±2.7

Table 2
Time-consuming of the fast "coarse-fine" registration method for segmentation of hippocampus (s).

Subjects	Resampling	ROI extraction	Non-rigid registration	Warp	Mean	Fuse (Our)	Total
No.1	4	1	3	0.5			
No.2	5	1	3	0.5			
No.3	4	1	2	0.5			
No.4	4	1	3	0.5	**8.88**	196	**267**
No.5	5	1	2	0.5			
No.6	5	1	4	0.5			
No.7	4	1	3	0.5			
No.8	5	1	3	0.5			

Table 3
Time-consuming of classical "coarse-fine" registration method for segmentation of hippocampus (s).

Subjects	Rigid registration	ROI extraction	Non-rigid registration	Warp	Mean
No.1	406	1	3	0.5	
No.2	745	1	3	0.5	
No.3	453	1	4	0.5	
No.4	450	1	3	0.5	575.13
No.5	765	1	3	0.5	
No.6	521	1	2	0.5	
No.7	517	1	3	0.5	
No.8	709	1	2	0.5	

3.2. Evaluation measures

Dice Similarity Coefficient, Jaccard Coefficient and Mean Distance (MD) are commonly applied to measure the performance of the segmentation approaches for hippocampus. They all measure the similarity between automatically segmented hippocampus and gold standard. In this paper, we also adopt these three evaluation metrics. Given the gold standard T and the automatic segmentation result S, the evaluation measures are defined as:

$$Dice(T, S) = \frac{2\|T \cap S\|}{\|T\| + \|S\|} \tag{20}$$

$$Jaccard(S, T) = \frac{VI(S \cap T)}{VI(S \cup T)} \tag{21}$$

$$MD = mean_{T \in BE}(\min_{S \in BF}\|S - T\|_2). \tag{22}$$

In above formulas, VI represents the volume size of the segmented region. BE and BF denote the boundary voxels of gold standard and segmented image respectively. The first two metrics measure the degree of overlap between automatically segmentation results and gold standard. The boundary difference is characterized by MD metric, which measures the surface distance between gold standard and automatically extracted results [45].

In this work, we apply the leave-one-out cross-validation method, in which one subject is selected as the target, and other images are used for atlas matching. Then, the differences between the automatically extracted results and the gold standard are calculated to verify the feasibility of our proposed method.

3.3. Parameter settings

There are mainly four parameters in our proposed method, including search radius r_s, patch radius r_p, the number of the most similar samples k in RDLF, and regularization parameter λ in RRDLF. In general, r_p is associated with the complexity of the structure, while r_s is related to the variability of the structure [46]. We investigated the effects of patch radius r_p and search radius r_s in the case of hippocampus segmentation. For example, the parameter settings proposed by Rousseau et al. [47] are that the patch radius is 1 voxels and the search radius is 5 voxels, whereas Platero

and Tobar [48] and Tong et al. [23] adopt a patch radius of 2 and a search radius of 3 voxels. In this paper, the patch radius is set to 2 and the search radius is tuned to 3. Regularization parameter λ is set to be $0.01\lambda_{max}$ referring to the penalty parameter λ in SPBM, where λ_{max} can be obtained automatically by the Lasso algorithm. And we set the number of the nearest neighbors $k = 9$ empirically.

Besides, our experiments use VS2010, ITK, CMAKE, VTK open source platforms, MATLAB and combine with ITK-SNAP to finally complete the left and right hippocampus segmentation under Windows 7 environment.

3.4. Fast "coarse-fine" registration performance

One of the major issues in multi-atlas segmentation methods is the computational burden of image registration. To reduce the high computational cost, a fast "coarse-fine" hybrid registration is proposed. Firstly, top 8 atlases are selected based on MI ranking between target image and atlas images. In order to describe succinctly and clearly, these selected atlases are renumbered from No.1 to No.8. After the atlas selection, each atlas image is registered to the target image using the proposed registration method and classical registration method, respectively. Tables 2 and 3 depict the time consuming for one randomly selected target image on IABA dataset. The first to fourth columns display the time spent in rigid registration or resampling operation, ROI extraction, non-rigid registration, and template warping process. The last column is the average time of the entire registration process for the selected 8 atlas images. As shown in Tables 2 and 3, the average time of the presented registration is only 8.88s, while the classical method takes 575.13s, which improves the speed of registration by more than 64 times. Table 2 also displays the time of label fusion and the entire inference time in the last two columns respectively, which shows the total time consumption of our segmentation framework is 267s.

For 10 randomly selected target images from IABA dataset, Table 4 shows the comparison segmentation results evaluated by metric Dice ratio using MV [19] method which is the simplest one. The results of the first line are obtained by the classical method. The second are generated based on our method. It is observed that the proposed "coarse-fine" registration method shows signif-
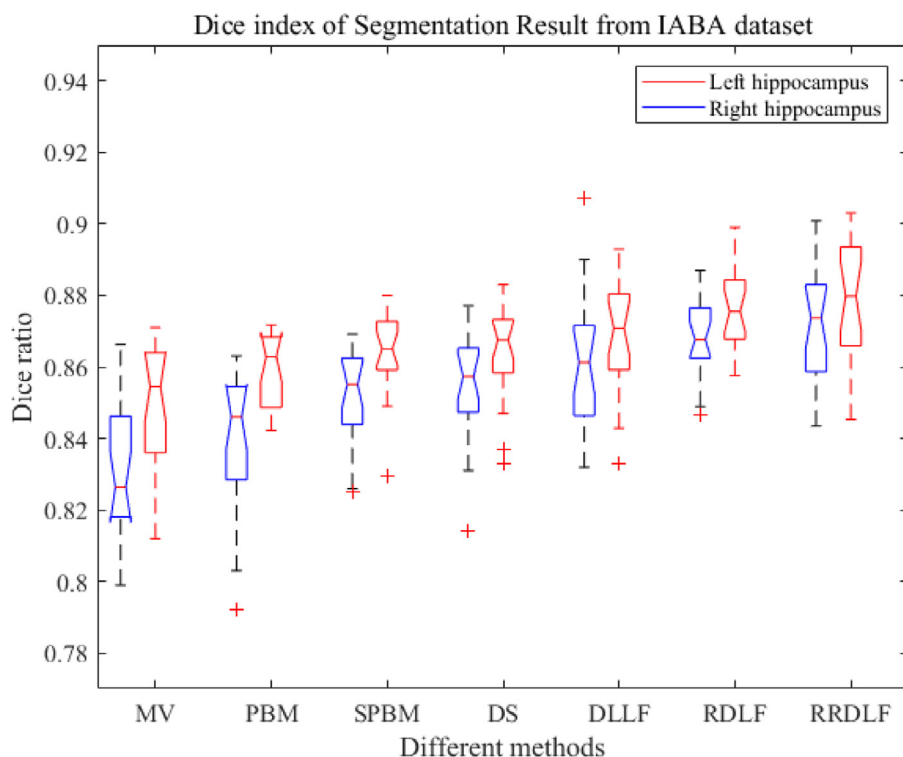
Fig. 7. Dice coefficient box-plot of the left and right hippocampus segmentation results of IABA dataset. Red box represents the left hippocampus segmentation result, while blue indicates right hippocampus segmentation result. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Dice ratio for our fast "coarse-fine" registration and classical "coarse-fine" registration (mean±std)

Registration method	Left hippocampus	Right hippocampus
Classical "coarse-fine" registration	0.83±0.033	0.85± 0.027
Fast "coarse-fine" registration	0.83± 0.034	0.86± 0.029

icant improvements in speed over classical "coarse-fine" registration without affecting the accuracy of segmentation.

3.5. Segmentation accuracy

To measure the performance of our proposed methods, we compare our method with MV [19], PBM [22], SPBM [23], DS [48] and DLLF. Among them, DS is a CRF model based on non-rigid registration method, DLLF is a convolutional method with deep learning using VGG-19. We denote the proposed robust discriminative label fusion algorithm RDLF, and denote RDLF with the segmentation refinement process in label space RRDLF.

3.5.1. Experimental results on IABA dataset
The box-plot shown in Fig. 7 illustrates the distribution of Dice for the proposed methods and the compared approaches on IABA dataset. For more intuitive observation, Fig. 8 shows the mean value of Dice ratio. As we can see, the proposed methods outperform other label fusion methods based on Dice metric, due to the robust discriminative observation given by our model.

To better compare the performance of segmenting both left and right hippocampus, Table 5 displays the values of Dice, Jaccard and MD obtained by different methods. From Table 5, the proposed method achieves the best performance with regard to three evaluation metrics. For instance, with respect to the Dice on the left hippocampus (the right hippocampus), RDLF improves 3.7%, 3.0%,

1.7%, 1.5%, and 0.9% (2.6%, 1.7%, 1.2%, 1.1%, and 0.7%) over MV, PBM, SPBM, DS and DLFF respectively. The average values of Jaccard and MD for segmenting left hippocampus (right hippocampus) are 0.765 and 0.243 (0.771 and 0.222) achieved by RDLF, which are higher than the second best results achieved by DLFF. As shown in Table 5, RRDLF increases the Dice values in segmenting the left and right hippocampus by 0.3% and 0.2% based on RDLF.

In addition to measuring the performance of the method via evaluation scores, Fig. 9 shows visual segmentation examples using different methods on hippocampus, where green denotes the gold standard, and red denotes the segmentation results obtained by different automatic labeling methods. It is known that the hippocampus is a relatively smooth organizational structure in biomedicine. As can be seen from these figures, the shape and size of the hippocampus extracted by the proposed method are very close to the gold standard, but smoother than the gold standard.

3.5.2. Experimental results on ADNI dataset
The performance of the proposed methods are also assessed on ADNI dataset. Figure 10 shows box-plot with the distribution of Dice for different methods in right and left hippocampus segmentation. Fig. 11 is a visual display of mean Dice average values of ADNI dataset. As we can see, compared with the conventional counterparts, the proposed methods have better performance in general. To better evaluate the performance, Table 6 shows the mean and standard deviation of Dice, Jaccard and MD across the target subjects.

As we can be seen from Table 6, the proposed method obtains the highest average Dice scores. RDLF is approximately 1.1% (0.7%) higher than suboptimal method DLLF for left hippocampus segmentation (right hippocampus segmentation). The average Jaccard and MD values generated by RDLF are 0.763 and 0.270 for left hippocampus (0.771 and 0.263 for right hippocampus), which is superior to each competitive approach (MV, PBM, SPBM, DS and
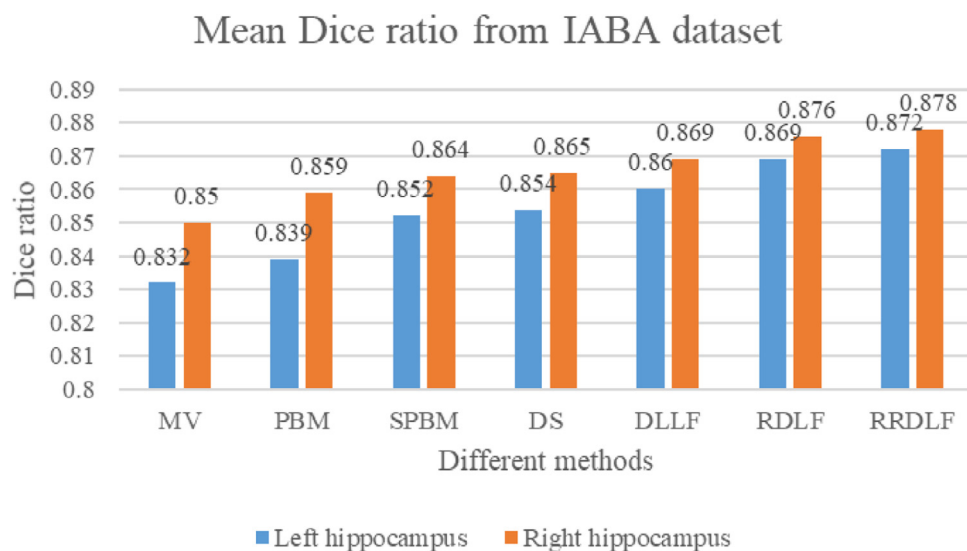
Fig. 8. Comparison of the average Dice values for segmenting left and right hippocampus using different methods from IABA dataset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5
The segmentation results of different methods using IABA dataset (mean±std).

Label fusion methods	Hippocampus	Dice	Jaccard	MD (mm)
MV	L	0.832±0.020	0.732±0.033	0.290±0.051
	R	0.850±0.017	0.747±0.029	0.263±0.047
PBM	L	0.839±0.020	0.736±0.030	0.260±0.044
	R	0.859±0.011	0.756±0.021	0.251±0.039
SPBM	L	0.852±0.014	0.748±0.027	0.247±0.043
	R	0.864±0.012	0.759±0.025	0.236±0.040
DS	L	0.854±0.015	0.752 ± 0.027	0.245 ± 0.037
	R	0.865 ± 0.013	0.761 ± 0.026	0.224 ± 0.038
DLFF	L	0.860 ± 0.033	0.756 ± 0.044	0.257 ± 0.048
	R	0.869 ± 0.016	0.767 ± 0.029	0.227 ± 0.043
RDLF	L	0.869 ± 0.011	0.765 ± 0.026	0.243 ± 0.038
	R	0.876 ± 0.012	0.771 ± 0.024	0.222 ± 0.036
RRDLF	L	**0.872 ± 0.016**	**0.772 ± 0.029**	**0.243 ± 0.040**
	R	**0.878 ± 0.016**	**0.774 ± 0.028**	**0.221 ± 0.039**

Table 6
The segmentation results of different methods using ADNI dataset (mean±std).

Label fusion methods	Hippocapus	Dice	Jaccard	MD (mm)
MV	L	0.854±0.016	0.742±0.029	0.325±0.050
	R	0.867±0.018	0.753±0.030	0.296±0.049
PBM	L	0.858±0.017	0743±0.028	0.314±0.042
	R	0.873±0.019	0.758±0.031	0.274±0.053
SPBM	L	0.864±0.022	0.747±0.035	0.283±0.048
	R	0.876±0.014	0.759±0.027	0.272±0.047
DS	L	0.868 ± 0.021	0.753 ± 0.033	0.280 ± 0.041
	R	0.878 ± 0.010	0.764 ± 0.023	0.265 ± 0.035
DLFF	L	0.868 ± 0.023	0.758 ± 0.034	0.278 ± 0.055
	R	0.88 ± 0.018	0.767 ± 0.029	0.269 ± 0.043
RDLF	L	0.879 ± 0.020	0.763 ± 0.02	0.270 ± 0.051
	R	0.887 ± 0.015	0.771 ± 0.027	0.263 ± 0.048
RRDLF	L	**0.882 ± 0.018**	**0.767 ± 0.029**	**0.268 ± 0.040**
	R	**0.889 ± 0.018**	**0.772 ± 0.029**	**0.262 ± 0.048**

DLFF). Generally, the proposed method RDLF significantly outperforms other methods in three evaluation metrics. Based on RDLF, RRDLF increases the Dice value by 0.2% to 0.3%, which is benefit from the full use of label information by SPBM method.

The visualization results for left and right hippocampus are shown in Fig. 12. The labeling results produced by the proposed methods are smoother than the gold standard. It is more biologically feasible for smooth results. However, the gold standard is manually labeled, there may be discontinuity errors between adja-cent slices. Therefore, the proposed segmentation method can preserve the shape and size of hippocampus well.

4. Discussion

In this work, a novel multi-atlas patch based label fusion method is developed for the segmentation of hippocampus. Most previously PBM methods select candidate patches adopting a predefined distance metric to measure the intensity-based similarity.
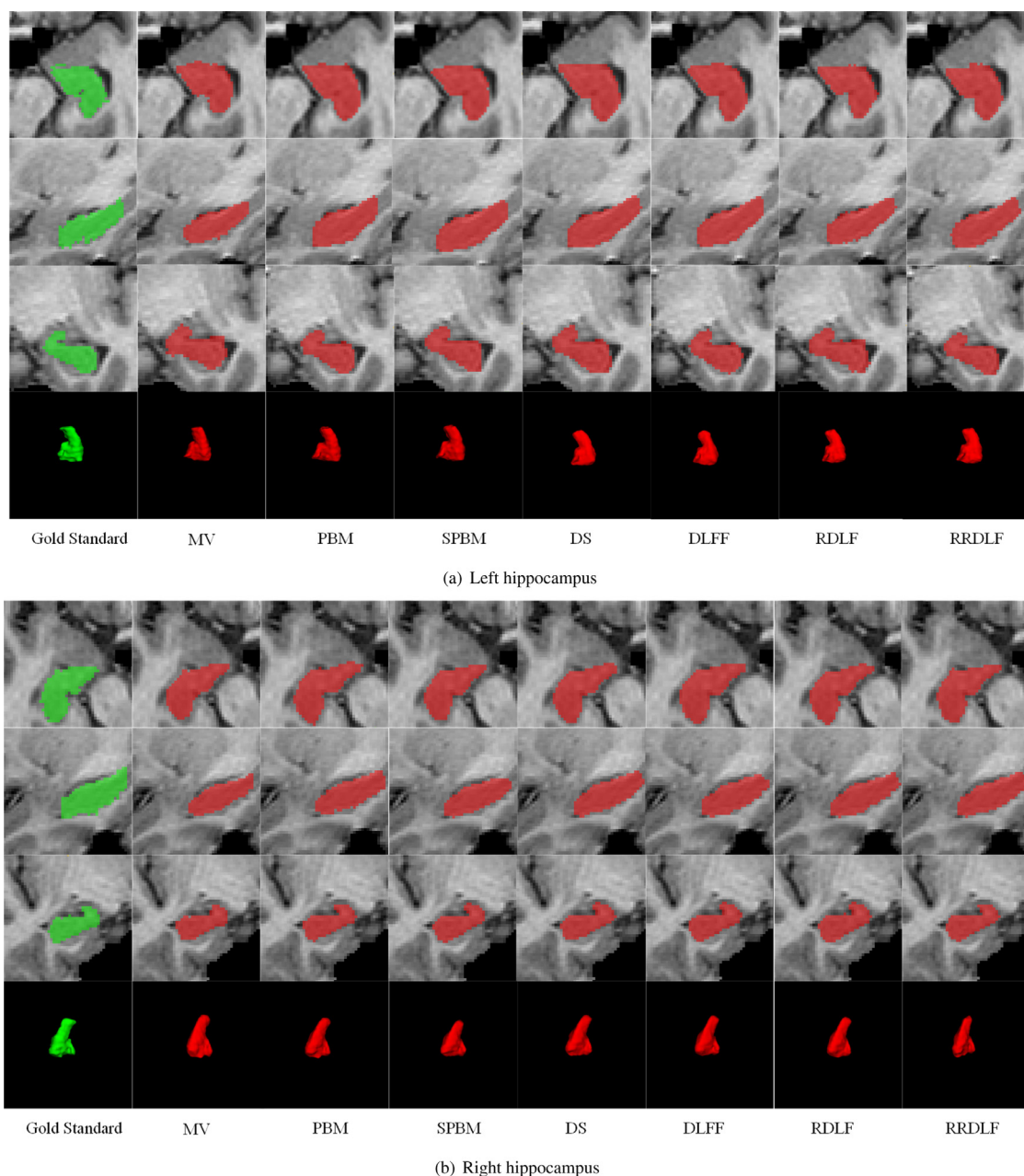
Fig. 9. Visual comparison of hippocampus segmentation results obtained by different methods on IABA dataset. Top to bottom: the slices display of the axial, sagittal, coronal and 3D visualization results. Left to right: gold standard, MV, PBM, SPBM, DS, DLFF, RDLF and RRDLF.

However, this kind of similarity measurement can not effectively characterize statistical distributions of image intensities. Besides most existing labeling fusion methods are typically based on the assumption that image patches with similar intensity share the same label. Nevertheless, this assumption is invalid in some situations, since similar patches may possess different labels. These make the segmentation results of hippocampus sensitive to the registration errors and the anatomical structure variability between different subjects.

To alleviate these drawbacks, the proposed method learns mapping functions between the patch and the label by robust distance metric learning. Meanwhile, with the spatial prior part, the pairwise homogeneity in classification labels and the pixel intensity is fully considered, thus to reduce the probability of misclassification. With the integrated formula of the objective function, the segmentation can be optimized by an effective graph cuts algorithm. As

we all know, the label information of atlases is rich of prior information and is very important for similarity measure. To take fully advantage of these prior knowledge, SPBM is used to refine the obtained segmentation results in label space. This approach turns out to be very efficient in estimating the final label of hippocampus.

In this study, the performance of the proposed method is compared with five multi-atlas based methods (MV, PBM, SPBM,DS, DLFF), in terms of Dice ratio, Jacard index and MD metrics on IABA and ADNI datasets. As shown in our experimental results, RDLF achieved the Dice of 86.9% and 87.6% on IABA dataset and Dice of 87.9% and 88.7% on ADNI dataset. The comparison results show that the proposed methods outperform current competitive methods in accordance with Dice, Jaccard and MD. Moreover, the labeling performance of RRDLF achieves a slight improvement on three evaluation metrics (e.g. the improvement of Dice value by 0.2% to 0.3%). The experiments on both IABA dataset and ADNI
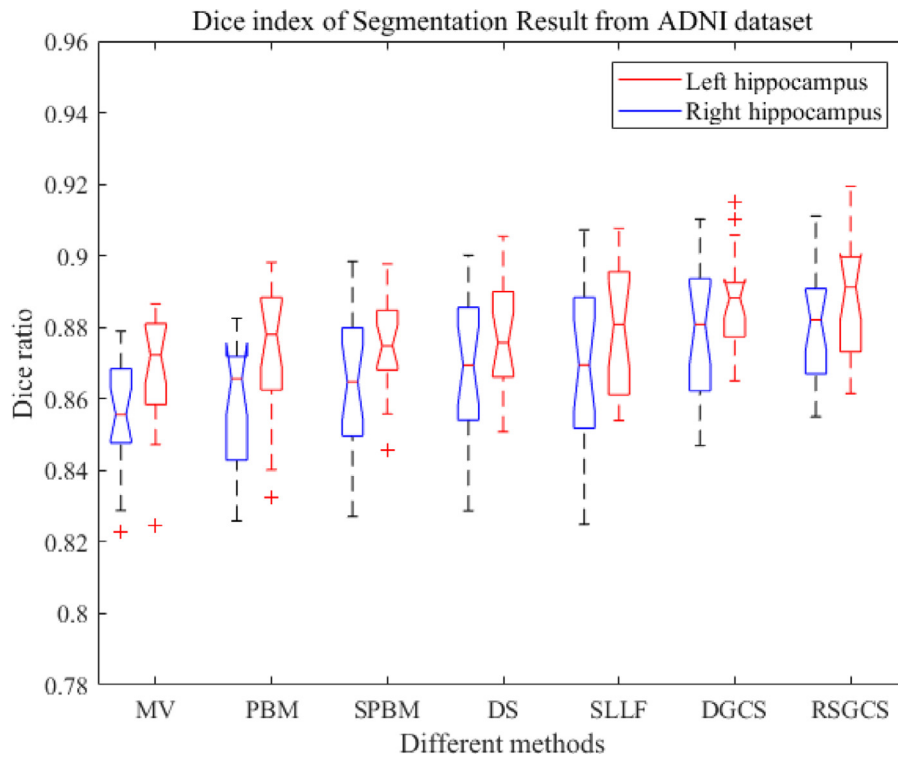
Fig. 10. Dice coefficient box-plot of the left and right hippocampus segmentation results of ADNI dataset. Red box represents the left hippocampus segmentation result, while blue indicates right hippocampus segmentation result. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
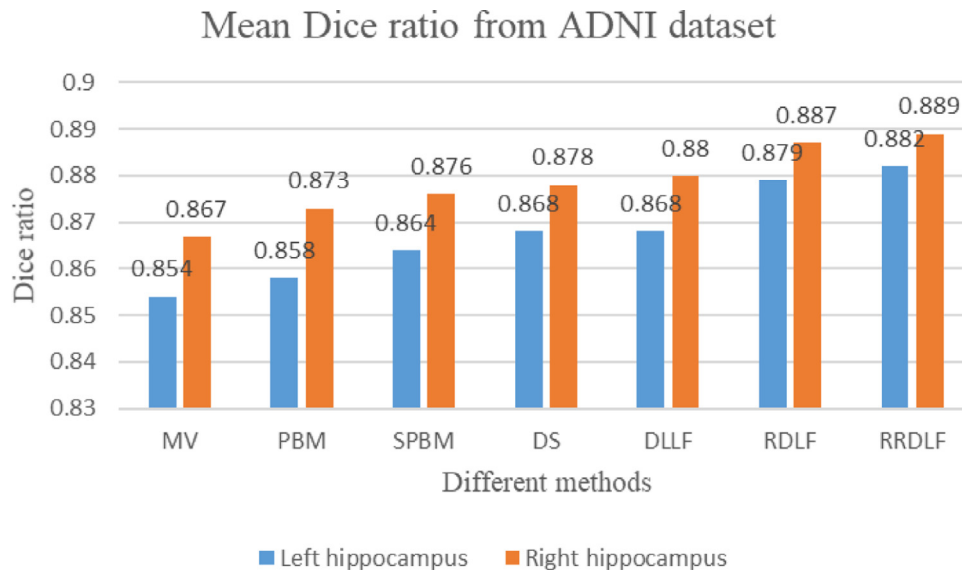

Fig. 11. Comparison of the average Dice values for segmenting left hippocampus and right hippocampus using different methods from ADNI dataset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

dataset show that the proposed method can acquire more promising performance for hippocampus segmentation. The experimental results demonstrate that RDLF significantly outperforms the competing methods. The reasons for the improvement are: (1) distance metric learning learns a more robust distance metric. It makes image patches which belongs to the same structure close to each other, while image patches of different structures are perfectly separated, and (2) spatial prior term fully considers the pairwise homogeneity to further improve classification accuracy. Experimental results also exhibit that RRDLF helps improve the labeling perfor-

mance for hippocampus segmentation. Since the target segmentation result is sparsely reconstructed in label space by RRDLF, so the label information is further used in overall labeling process.

In addition, one of the main issues in the multi-atlas segmentation is the computational cost for image registration. To alleviate this drawback, we also proposed a fast "coarse-fine" hybrid registration at the same time. Since resampling does not require any iterative optimization process compared with rigid registration, as a result, the computational burden is reduced. Experimental results show it improves the speed of registration by more than 64
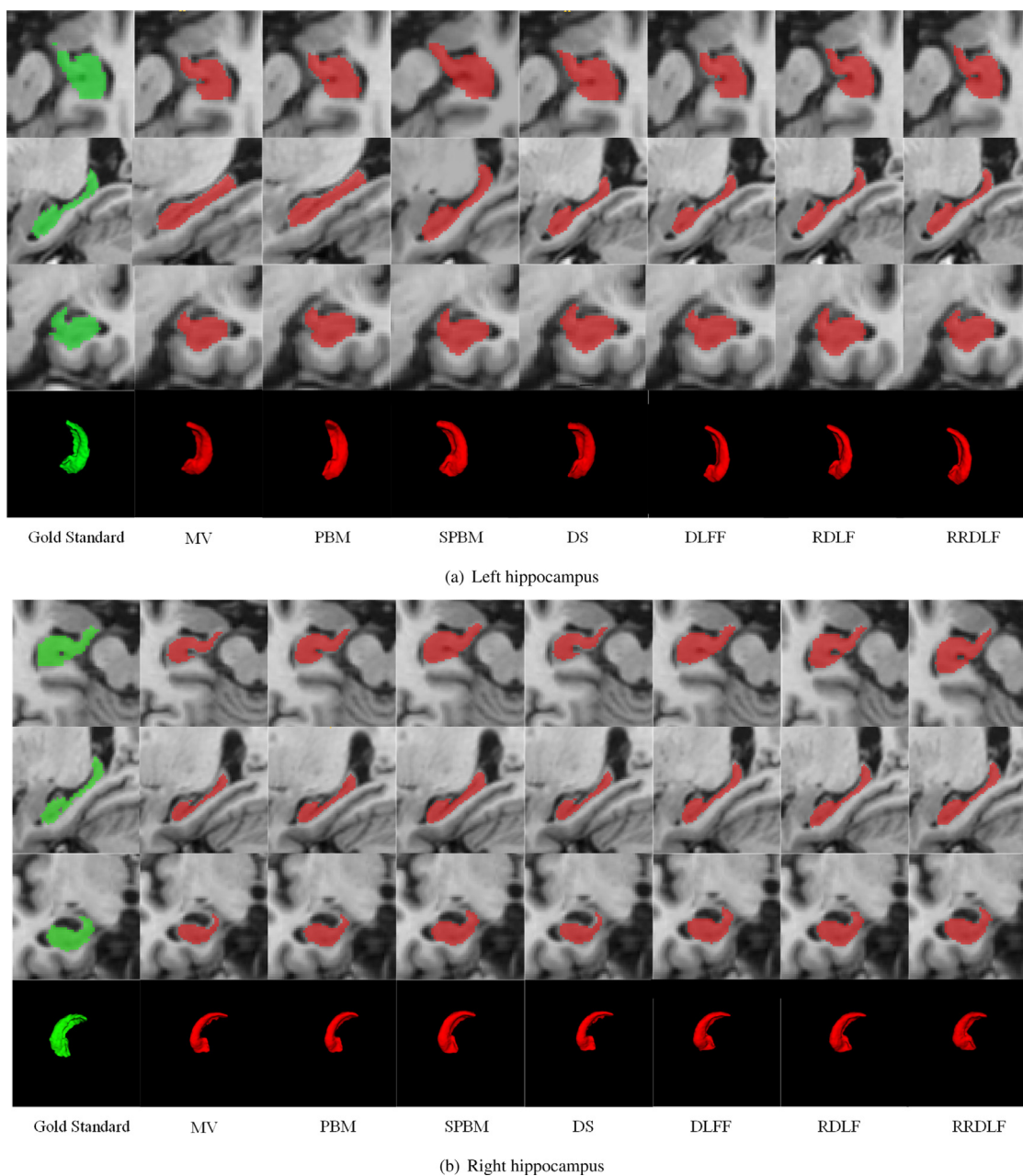
Fig. 12. Visual comparison of hippocampus segmentation results obtained by different methods on ADNI dataset. Top to bottom: the slices display of the axial, sagittal, coronal and 3D visualization results. Left to right: gold standard, MV, PBM, SPBM, DS, DLFF, RDLF and RRDLF.

times compared to classical "coarse-fine" hybrid registration, which demonstrates the feasibility of this strategy.

5. Conclusions

In this paper, an automatic labeling method is developed to segment the hippocampus from brain MR images. It is a patch embedding multi-atlas label fusion method based on CRF model that combines distance metric learning and graph cuts. The proposed label fusion strategy is evaluated on IABA dataset and ADNI dataset. Experimental results show that the proposed algorithm outperforms other competing hippocampus segmentation methods. Meanwhile, to avoid the high computational burden of image registration, a fast "coarse-fine" hybrid registration method is proposed, which can improved the speed of registration effectively. We hope that the proposed model can be developed into a gen-

eral model, which can easily combine machine learning and deep learning algorithms to segment the hippocampus. In the future work, we are interested in developing more robust statistical label fusion strategy, and applying the proposed framework to automatically classify patients with AD and/or MCI from hippocampus segmentations extracted in brain MR image to better satisfy the clinical needs.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgement

This research is supported by the National Natural Science Foundation of China (Grant Nos. 61971356, 61801395, 61801393

References

[1] Y. Shachor, H. Greenspan, J. Goldberger, A mixture of views network with applications to multi-view medical imaging, Neurocomputing 374 (2020) 1–9.

[2] J. Li, Z.L. Yu, Z. Gu, H. Liu, Y. Li, Mman: Multi-modality aggregation network for brain segmentation from mr images, Neurocomputing 358 (2019) 10–19.

[3] H. Jia, Y. Xia, Y. Song, W. Cai, M. Fulham, D.D. Feng, Atlas registration and ensemble deep convolutional neural network-based prostate segmentation using magnetic resonance imaging, Neurocomputing 275 (2018) 1358–1369.

[4] N. Noorizadeh, K. Kazemi, H. Danyali, A. Aarabi, Multi-atlas based neonatal brain extraction using a two-level patch-based label fusion strategy, Biomed. Signal Process. Control 54 (2019) 101602.

[5] L. Sun, W. Shao, M. Wang, D. Zhang, M. Liu, High-order feature learning for multi-atlas based label fusion: application to brain segmentation with MRI, IEEE Trans. Image Process. (2019).

[6] L. Sun, C. Zu, W. Shao, J. Guang, D. Zhang, M. Liu, Reliability-based robust multi-atlas label fusion for brain MRI segmentation, Artif. Intell. Med. 96 (2019) 12–24.

[7] H. Zhu, Z. Tang, H. Cheng, Y. Wu, Y. Fan, Multi-atlas label fusion with random local binary pattern features: application to hippocampus segmentation, Sci. Rep. 9 (1) (2019) 1–14.

[8] D. Cárdenas-Peña, A. Tobar-Rodríguez, G. Castellanos-Domínguez, A.D.N. Initiative, et al., Adaptive Bayesian label fusion using kernel-based similarity metrics in hippocampus segmentation, J. Med. Imaging 6 (1) (2019) 014003.

[9] H. Zhu, G. He, Joint neighboring coding with a low-rank constraint for multi-atlas based image segmentation, J. Med. Imaging Health Inf. 10 (2) (2020) 310–315.

[10] J. Wu, X. Tang, Brain segmentation based on multi-atlas guided 3D fully convolutional network ensembles, arXiv preprint arXiv:1901.01381(2019).

[11] Y. Guo, Z. Wu, D. Shen, Learning longitudinal classification-regression model for infant hippocampus segmentation, Neurocomputing (2019).

[12] Yang, Medical image computing and computer-assisted intervention, Lect. Notes Comput. Sci. 6361 (1) (2009) 69–87.

[13] L. Zhang, Q. Wang, Y. Gao, H. Li, G. Wu, D. Shen, Concatenated spatially-localized random forests for hippocampus labeling in adult and infant MR brain images, Neurocomputing 229 (2017) 3–12.

[14] O.M. Benkarim, G. Piella, N. Hahner, E. Eixarch, M.A.G. Ballester, G. Sanroma, Patch spaces and fusion strategies in patch-based label fusion, Comput. Med. Imaging Graphics 71 (2019) 79–89.

[15] Q. Zheng, Y. Fan, Integrating semi-supervised label propagation and random forests for multi-atlas based hippocampus segmentation, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, 2018, pp. 154–157.

[16] J. Alvén, F. Kahl, M. Landgren, V. Larsson, J. Ulén, O. Enqvist, Shape-aware label fusion for multi-atlas frameworks, Pattern Recognit. Lett. 124 (2019) 109–117.

[17] H.W. Lee, M.R. Sabuncu, A.V. Dalca, Few labeled atlases are necessary for deep-learning-based segmentation, arXiv preprint arXiv:1908.04466(2019).

[18] X. Lin, X. Li, Image based brain segmentation: from multi-atlas fusion to deep learning, Curr. Med. Imaging Rev. 15 (5) (2019) 443–452.

[19] R.A. Heckemann, J.V. Hajnal, P. Aljabar, D. Rueckert, A. Hammers, Automatic anatomical brain MRI segmentation combining label propagation and decision fusion, NeuroImage 33 (1) (2006) 115–126.

[20] S.K. Warfield, K.H. Zou, W.M. Wells, Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation, IEEE Trans. Med. Imaging 23 (7) (2004) 903–921.

[21] G. Sanroma, O.M. Benkarim, G. Piella, O. Camara, G. Wu, D. Shen, J.D. Gispert, J.L. Molinuevo, M.A.G. Ballester, A.D.N. Initiative, et al., Learning non-linear patch embeddings with neural networks for label fusion, Med. Image Anal. 44 (2018) 143–155.

[22] P. Coupé, J.V. Manjón, V. Fonov, J. Pruessner, M. Robles, D.L. Collins, Nonlocal patch-based label fusion for hippocampus segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2010, pp. 129–136.

[23] T. Tong, R. Wolz, P. Coupé, J.V. Hajnal, D. Rueckert, A.D.N. Initiative, et al., Segmentation of mr images via discriminative dictionary learning and sparse coding: application to hippocampus labeling, NeuroImage 76 (2013) 11–23.

[24] L. Fang, L. Zhang, D. Nie, X. Cao, I. Rekik, S.-W. Lee, H. He, D. Shen, Automatic brain labeling via multi-atlas guided fully convolutional networks, Med. Image Anal. 51 (2019) 157–168.

[25] Y. Liu, Y. Wei, C. Wang, Subcortical brain segmentation based on atlas registration and linearized kernel sparse representative classifier, IEEE Access 7 (2019) 31547–31557.

[26] H. Yang, J. Sun, H. Li, L. Wang, Z. Xu, Deep fusion net for multi-atlas segmentation: application to cardiac mr images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 521–528.

[27] P. Aljabar, R.A. Heckemann, A. Hammers, J.V. Hajnal, D. Rueckert, Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy, NeuroImage 46 (3) (2009) 726–738.

[28] G. Sanroma, G. Wu, Y. Gao, D. Shen, Learning to rank atlases for multiple-atlas segmentation, IEEE Trans. Med. Imaging 33 (10) (2014) 1939–1953.

[29] H. Zhu, H. Cheng, X. Yang, Y. Fan, A.D.N. Initiative, et al., Metric learning for multi-atlas based segmentation of hippocampus, Neuroinformatics 15 (1) (2017) 41–50.

[30] H. Wang, J.W. Suh, S.R. Das, J.B. Pluta, C. Craige, P.A. Yushkevich, Multi-atlas segmentation with joint label fusion, IEEE Trans. Pattern Anal. Mach.Intell. 35 (3) (2012) 611–623.

[31] H. Liu, M. Yan, E. Song, Y. Qian, X. Xu, R. Jin, L. Jin, C.-C. Hung, Label fusion method based on sparse patch representation for the brain MRI image segmentation, IET Image Process. 11 (7) (2017) 502–511.

[32] T. Vercauteren, X. Pennec, A. Perchant, N. Ayache, Diffeomorphic demons: efficient non-parametric image registration, NeuroImage 45 (1) (2009) S61–S72.

[33] L. Yang, R. Jin, R. Sukthankar, Bayesian active distance metric learning, 2012, arXiv preprint arXiv:1206.5283

[34] K.Q. Weinberger, L.K. Saul, Distance metric learning for large margin nearest neighbor classification, J. Mach. Learn. Res. 10 (Feb) (2009) 207–244.

[35] M.J. Wainwright, M.I. Jordan, et al., Graphical models, exponential families, and variational inference, Found. Trends® Mach.Learn. 1 (1–2) (2008) 1–305.

[36] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT press, 2016.

[37] D. Wang, X. Tan, Robust distance metric learning via Bayesian inference, IEEE Trans. Image Process. PP (99) (2017). 1–1

[38] M.D. Hoffman, D.M. Blei, Structured stochastic variational inference, Comput. Sci. 14 (1) (2014) 1303–1347.

[39] Y. Boykov, G. Funka-Lea, Graph cuts and efficient nd image segmentation, Int. J. Comput. Vis. 70 (2) (2006) 109–131.

[40] M. Wels, G. Carneiro, A. Aplas, M. Huber, J. Hornegger, D. Comaniciu, A discriminative model-constrained graph cuts approach to fully automated pediatric brain tumor segmentation in 3-D MRI, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2008, pp. 67–75.

[41] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Trans. Pattern Anal. Mach.Intell. 31 (2) (2008) 210–227.

[42] F. Zhang, Z. Li, B. Zhang, H. Du, B. Wang, X. Zhang, Multi-modal deep learning model for auxiliary diagnosis of Alzheimer's disease, Neurocomputing 361 (2019) 185–195.

[43] M. Liu, J. Zhang, P.-T. Yap, D. Shen, View-aligned hypergraph learning for Alzheimer's disease diagnosis with incomplete multi-modality data, Med. Image Anal. 36 (2017) 123–134.

[44] C.R. Jack Jr, M.A. Bernstein, N.C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P.J. Britson, J. L. Whitwell, C. Ward, et al., The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods, J. Magn. Reson. Imaging 27 (4) (2008) 685–691.

[45] Y. Hao, T. Wang, X. Zhang, Y. Duan, C. Yu, T. Jiang, Y. Fan, A.D.N. Initiative, Local label learning (LLL) for subcortical structure segmentation: application to hippocampus segmentation, Hum. Brain Mapp. 35 (6) (2014) 2674–2697.

[46] P. Coupé, J.V. Manjón, V. Fonov, J. Pruessner, M. Robles, D.L. Collins, Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation, NeuroImage 54 (2) (2011) 940–954.

[47] F. Rousseau, P.A. Habas, C. Studholme, A supervised patch-based approach for human brain labeling, IEEE Trans. Med. Imaging 30 (10) (2011) 1852–1862.

[48] C. Platero, M.C. Tobar, Combining a patch-based approach with a non-rigid registration-based label fusion method for the hippocampal segmentation in Alzheimer's disease, Neuroinformatics 15 (2) (2017) 165–183.