



# White Matter Segmentation Algorithm for DTI Images Based on Super-Pixel Full Convolutional Network

Yiping Mu<sup>1</sup> · Qi Li<sup>1</sup> · Yang Zhang<sup>1</sup>

Received: 20 May 2019 / Accepted: 29 July 2019 / Published online: 12 August 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

Diffusion tensor imaging (DTI) is a new imaging method that can be used to non-invasively measure the diffusion coefficient of water molecules in biological tissue structures in recent years. Since the DTI data is a tensor space, its segmentation is different from ordinary MRI images. Based on the existing deep learning model, an improved image semantic segmentation method based on super-pixels and conditional random field is proposed. Firstly, this paper uses the existing feature extraction model based on deep learning to obtain rough semantic segmentation results, including high-level semantic information of the image but lacking details of the image. In addition, the super-pixel segmentation algorithm is implemented to obtain super-pixels that carries more low-level information. Secondly, due to the lack of image details in rough segmentation results, the segmentation of the edge of the image is inaccurate. In this paper, a boundary optimization algorithm is proposed to optimize the edge segmentation accuracy of the rough results. Finally, the use of super-pixels for local boundary optimization can improve the segmentation accuracy. Experiments results show that this segment is a practical and effective method.

**Keywords** Diffusion tensor imaging · Rough segmentation · Boundary optimization · Super-pixel segmentation · Full convolutional network · Deep learning

## Introduction

Diffusion tensor imaging (DTI) is of great significance and clinical importance in the study of cognitive function and neural activity of brain [1]. DTI gets self-diffusion tensor images by measuring the diffusing characteristics of water molecules in brain tissues based on diffusion weighted imaging (DWI) [2]. Compared to T2-weighted MRI, DTI could get the information of the asymmetric construction distribution of brain tissues [3].

Image segmentation is an important step and a key link in medical image analysis, understanding and description [4]. The head model can be divided into five types of tissues: scalp, skull, gray matter, white matter, and cerebrospinal fluid. The physiological functions and structures of each tissue are

different [5]. Different brain images have different imaging principles and different segmentation methods [6]. Brain tissue segmentation is a key component in both electroencephalogram/magnetic science research and clinical diagnosis of brain diseases. The distribution of internal tissues of the brain is very complicated, and different brain tissues are intertwined. If the different brain tissues cannot be accurately segmented, the correct conductivity of the brain tissue cannot be assigned, and finally accurate calculation results cannot be obtained [7].

Diffusion tensor imaging (DTI) is a new imaging method that can be used to non-invasively measure the diffusion coefficient of water molecules in biological tissue structures in recent years. This technique has a good application prospect in the non-invasive tracking of white matter fiber bundles, which can further help detect abnormalities of early white matter lesions, and has superiority that can not be compared with general MRI examination. Since the DTI data is a tensor space, its segmentation is different from ordinary MRI images. Firstly, the DWI data is converted into DTI data according to the magnetic field gradient table and the  $b$  value, and then divided according to the diffusion characteristics of different tissue water molecules.

---

This article is part of the Topical Collection on *Image & Signal Processing*

✉ Yiping Mu  
yiping.mu@symc.edu.cn

<sup>1</sup> Central Hospital Affiliated of Shenyang Medical College, Shenyang 110024, Liaoning, China

Recently, there are three kinds of segmentation methods based on diffusion tensor imaging. Literature [8] proposed support vector machine and Markov random model segmentation algorithm; since support vector machine training is difficult to converge, it is difficult to choose appropriate tensor training sample set; the hidden Markov random model and the maximum expectation algorithm is proposed by literature [9], where the segmentation error is large and the efficiency is low; literature [10] proposed statistical fuzzy parameterless segmentation, but its error is very large without priori knowledge combining different brain tissue diffusion characteristics. In order to improve the accuracy of segmentation, domestic and foreign scholars have proposed many excellent segmentation algorithms.

An improved fuzzy C mean clustering algorithm is proposed in literature [11], the fuzzy C-means clustering algorithm randomly selects the center point of the initial cluster. Because the final clustering result has a certain dependence on the initial clustering center point, the random clustering center point will not affect the final clustering. As a result, the combination of FCM and maximum-minimum distance algorithm is proposed. The maximum and minimum distances are introduced on the basis of classical FCM, and the improved fuzzy C-means clustering algorithm is validated by the experimental data set. The experimental results show that the improved fuzzy C-means segmentation can obtain more smooth edge information, the wrong segmentation area is reduced, and the segmentation accuracy is improved. An adaptive mean drift algorithm is proposed in literature [12], the basic idea of the mean shift algorithm is to search for the densest area of sample points in the given sample space and drift to the local density maximum along the direction of increasing density. Unlike other clustering algorithms, drifting to find the local maximum is a continuous iterative process, so no prior knowledge is required. However, the bandwidth of the traditional mean shift algorithm is a fixed value and cannot be automatically adjusted according to the distribution of the pixel points. Therefore, their improved algorithm proposes an adaptive mean shift algorithm. By redefining the window function and combining the probability density function of the pixels, the pixel points are different. Probability density applies to different bandwidth values, and the segmentation effect is improved in DTI image segmentation.

In view of the brain white matter fibers segmentation from diffusion tensor image, literature [13] presented an algorithm based on Riemannian manifold. Firstly, construct a  $3 \times 3$  symmetric positive definite covariant tensor for each voxel using diffusion tensor image, by which tensor field is constructed to illuminate brain white matter. Secondly, the tensor field is regarded as Riemannian

manifold, and the fluid motion in the tensor field is expressed by Navier-Stoke equation, so the problem of brain white matter fibers between two voxels can be transformed into the computation of the smallest distance between two points in Riemannian manifold. Finally, distances between two points in Riemannian manifold [14], which are the brain white matter fibers, can be expressed by geodesic, whose numerical solution is based on Level-Set algorithm. Compared with the conventional brain white matter fibers segmentation, this algorithm's accuracy and robustness are greatly improved.

However, the DTI segmentation algorithms mentioned above are all based on low-level learning methods such as traditional image processing and pattern recognition [15]. They are sensitive to noise and can not effectively process and analyze large-scale image data [16]. Compared with the traditional shallow learning, the deep learning represented by convolutional neural network emphasizes the depth of model structure and the importance of feature learning. By layer-by-layer feature transformation, the feature representation of samples in the original space is transformed into a new deep-feature space, which makes classification or prediction easier. Compared with the method of constructing features by artificial rules, using large data to learn features can better represent the rich intrinsic information of data. Shallow learning lacks generalization ability for complex problems due to the limitations of samples and computational capacity [17]. Deep learning can achieve complex function approximation by learning a deep non-linear network structure, demonstrating a powerful ability to learn the essential characteristics of data from the sample set.

Based on the existing deep learning model, an improved image semantic segmentation method based on super-pixels and conditional random fields is proposed. Firstly, this paper uses the existing feature extraction model based on deep learning to obtain rough semantic segmentation results, including high-level semantic information of the image but lacking details of the image. In addition, the super-pixel segmentation algorithm is applied to obtain super-pixels that carries more low-level information. Secondly, due to the lack of image details in rough segmentation results, the segmentation of the edge of the image is inaccurate. In this paper, a boundary optimization algorithm is proposed to optimize the edge segmentation accuracy of the rough results. The edge segmentation effect in the rough results is preliminarily optimized using super-pixels. Finally, the use of super-pixels for local boundary optimization can improve the segmentation accuracy. In order to further improve the segmentation accuracy, the fully connected conditional random field is used to constrain the pixels with similar structure and spatial position, and make full use of the local texture features, global context information and smooth priors to further optimize the semantic segmentation results of the image.

## Materials and methods

### Strategy

The algorithm proposed in this paper is essentially a two-stage segmentation algorithm [18]. The main differences between it and the traditional image segmentation algorithm are as follows: (1) the traditional image segmentation algorithm usually is prone to over-segmentation first, and then uses different classification models to semantically classify the over-segmented regions. However, the proposed algorithm does not need to preprocess the image; (2) the front-end classification model FCN used in this paper is a full-convolution network. The traditional image segmentation method based on convolution neural network adopts the region-based R-CNN model. The two networks are essentially different in structure and function. The FCN model implements the classification at the pixel level whose classification label is the label of a single pixel. And the R-CNN model implements the classification at the regional level whose classification label is the label of the smallest rectangular area including the object in the image [19]. The classification information provided by the FCN model is more comprehensive and detailed, while the R-CNN model can only limit the classification precision to the over-segmented region, and the classification precision of the former is higher; (3) the post-processing step in two-stage model adopted in this paper is a progressive optimization algorithm. Different from the traditional optimization algorithm based on probability graph model, this paper makes full use of the good consistency of super-pixels to the object edge, which not only optimizes the output of the front-end FCN model, but also plays a better reference role in constructing the following conditional random field model.

### Extraction of rough features

In this paper, FCN model is used to extract rough features in images. Different from the traditional convolution neural network model, FCN model can input any size of image and generate corresponding size of output to obtain pixel-level classification results [20]. The FCN model can be transformed from the existing convolutional neural network. And the FCN model used in this paper is transformed from VGG-16 in series VGGNet.

There are six kinds of network structures in VGGNet series. As shown in Table 1, there are some differences in the structure and configuration of each network [21]. Among them, different letters denote different types of networks. VGG-16 is one of the network D. The range of convolution layers of these six networks is 11–19 layers. Table 2 shows the number of network parameters. It can be

seen from Table 2 that although the depth of network A to network E increases gradually, the number of network parameters does not increase with it, because a large number of parameters are used in the last three full-connection layers. Although the number of parameters of convolution layer is not large, the calculation of convolution operation is pretty large, which leads to the big time cost in convolution operation for the training process.

Compared with other deep convolution neural networks, VGG-16 network can achieve a better balance between feature extraction and training efficiency [22]. GooleNet and ResNet have deeper layers and the ability to extract more abstract image features. However, because of too many layers, deep networks are prone to over-fitting and gradient dispersion when training large-scale data sets, which affects the segmentation effect. Because VGG-16 network uses smaller  $3 \times 3$  convolution core, it improves the non-linear fitting ability of the model, and to a certain extent, it makes up for the deficiency of feature extraction ability caused by fewer network layers. According to the research results proposed by literature [23], through testing on multiple data sets, the top-1 error rate of VGG-16 and VGG-19 is almost the same, while the top-5 error rate is only 0.1% gap. This shows that VGG-19 does not significantly improve the segmentation precision, so it is better to use VGG-16 with fewer network layers while guaranteeing the ability of feature extraction. In contrast, AlexNet has too few network layers to meet the requirement of image segmentation for feature extraction. Considering the feature extraction ability of the network and the resource occupancy rate of network training, our proposed model transforms VGG-16 network to construct the FCN model.

The specific method of converting VGG-16 network into FCN model is to replace the full connection layer of the original network with convolution layer, and retain the first five layers of structure [24]. In the whole process of feature extraction, the resolution of feature mapping is getting lower and lower after several iterations of convolution and pooling. In order to reconstruct the final output as the input image that both of them are having the same size, an up-sampling operation is needed for the intermediate output. Compared with the original input image, the resolution of the final output feature mapping is reduced by 2, 4, 8, 16 and 32 times. The results of FCN-32 s can be obtained by directly sampling the rough features of the last layer 32 times. However, the output image of FCN-32 s lacks many details because of the large magnification, so its results are not accurate enough. In order to improve the accuracy, we need to add more details in the next several layers to the FCN-32 s. By combining more details with the output of FCN-32 s, we can further obtain the results of FCN-16 s and FCN-8 s. The effect of FCN-32 s, FCN-16 s and FCN-8 s are tested in this paper. Figure 1 shows the image segmentation results under three receptive fields.

**Table 1** VGGNet structure configuration table

Network layer VGGNet	A(11)	A-LRN(11)	B(13)	C(16)	D(16)	E(19)
Input layer	320 × 240					
Convolution layer	conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
Pooling layer	maxpool					
Convolution layer	conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
Pooling layer	maxpool					
Convolution layer	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 Conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
Pooling layer	maxpool					
Convolution layer	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 Conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
Pooling layer	maxpool					
Convolution layer	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 Conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
Pooling layer	maxpool					
FC layer	FC-4096					
FC layer	FC-4096					
FC layer	FC-100					
Classification layer	softmax					

An important problem to be solved in image segmentation is how to combine target recognition and target positioning organically. In other words, segmentation is to classify pixels by pixels and merge location and classification information. On the one hand, because of the difference of receptive field, the resolution is relatively high after the previous convolution operations, and the classification of pixels is not accurate, but the positioning of each pixel is comparatively accurate. On the other hand, the resolution is relatively low in the last several convolutions and the positioning of the pixels is not accurate enough, but the classification of the pixels is comparatively accurate. As can be seen from Fig. 1, the result of FCN-32 s is relatively too smooth but the edge segmentation is not very accurate. This is because the receptive field of FCN-32 s model is larger and more suitable for macro-perception [25]. In contrast, the receptive field of FCN-8 s model is smaller and more suitable for details. As can be seen from Fig. 1, the results of FCN-8 s are the closest to the true semantic labels, and are significantly better than those of FCN-16 s and FCN-32 s. Therefore, FCN-8 s is used as the front-end to extract the rough features of the image. Nevertheless, the results of FCN-8 s are still insensitive to the details of the image. Therefore, two-level optimization algorithm is used for fine segmentation.

## Fine segmentation based on two-level optimization

Since the most important characteristic of super-pixel is that it can fit the image edge, we propose to use super-pixel to optimize the rough features extracted from the front-end so as to improve the segmentation accuracy of image edge [26]. Usually, a super-pixel can be regarded as a set of similar pixels such as position, color, texture, etc. Although the super-pixel is still a set of pixels according to this similarity, compared with the pixel, the super-pixel has certain visual significance. Although a single super-pixel does not have effective semantic information, a single super-pixel is an object or a part of an object with semantic information.

## Framework of edge optimization algorithm

Figure 2 illustrates the flow chart of using super-pixels to optimize the effect of segmentation on object edges. The core idea is to use the Simple Linear Iterative Clustering super-pixel segmentation algorithm to generate super-pixels, and then optimize rough features through the object edges that well fitted by super-pixels. This optimization algorithm can improve the segmentation accuracy of object edges. It is

**Table 2** The number of arguments for different blocks

Blocks	A,A-LRN	B	C	D	E
Arguments	133	133	134	138	145

divided into three steps. Firstly, the rough features of input image are extracted and semantically labeled by using the FCN model mentioned above. Secondly, the Simple Linear Iterative Clustering super-pixel segmentation algorithm is used to segment the input image and obtain the super-pixel segmentation image. Thirdly, combining the two kinds of segmentation images obtained in the first two steps, the proposed optimization algorithm is used to optimize the local edges of the rough features obtained in the first step. After edge optimization, the segmentation results will have more accurate edge information than the results of FCN model, and also reduce the diffusion error in the process of up-sampling. These two points make the accuracy of segmentation improved.

**Super-pixel-based edge optimization algorithm**

The key of local edge optimization is the effect of combining super-pixel edge with FCN model. If the effect is good, it will improve the overall effect [27]. On the contrary, it will not improve the segmentation effect, but will expand the segmentation error and reduce the accuracy of segmentation. The optimization algorithm proposed in this paper is shown in Algorithm 1, and the core idea of edge optimization is shown in pseudo-code. The core idea of the edge optimization algorithm proposed in this paper is to semantically assign labels to all the pixels in the super-pixel using the pixel-level characteristic pattern output by FCN. There are several possible situations in this process. According to whether or not the image edge is included in the super-pixel, there are two kinds of cases, namely, edge and no edge. In the case of super-pixels containing edges, there are two cases according to whether all pixels have the same semantic label. For the convenience of description, case A is that a single super-pixel does not contain image edges, and all the pixels in the super-pixel have the same semantic label; while case B is that the super-pixel does not contain image edges, but the pixels in the super-pixel have multiple semantic labels. Case C is that the super-pixel contains image edges, but all pixels still have the same semantic

label; case D is that the super-pixel contains image edges and the pixel has multiple semantic labels.

**Fine segmentation**

After local edge optimization, it is still necessary to improve the segmentation accuracy of weak edges, small structures and complex scenes. Therefore, this paper uses the fully connected conditional random field model to restore the edge of the image more accurately, that is, to further optimize the segmentation effect of the edge part of the image, and then to improve the overall accuracy of image segmentation.

A fully connected conditional random field model is established by using image  $I$  and edge-optimized pixel-level segmented graph, and it is represented by probability distribution  $P(X)$ . The energy function used in this paper is  $E(x)$ , in which one-dimensional potential energy  $\psi_u(x_i)$  is the optimized segmentation results of local edges and represents the probability that the pixel  $i$  is labeled as label  $x_i$ . But the fully connected conditional random field has a large number of edges, which makes the calculation very difficult. For example, Robust CRF algorithm takes several hours to process a picture. So this paper uses mean field approximation and high-dimensional filtering to solve the probability distribution  $P(X)$  of CRF efficiently.

**Mean field approximation**

Since calculating accurate probability distribution  $P(X)$  is a NP-Hard problem, a more general approach is to find probability distribution  $Q(X)$  from distribution set  $Q$  by minimizing *KL (Kullback-Leibler) divergence to approximately replace probability distribution  $P(X)$ . Probability distribution set  $Q$  can be expressed as the product of independent edge probability distribution, as shown in formula (1):*

$$Q(X) = \prod Q_i(x_i) \tag{1}$$

where  $Q_i(x_i)$  is the edge probability distribution of variable  $x_i$ . And  $Q_i(x_i)$  can be obtained by using KL divergence and Lagrange equation.

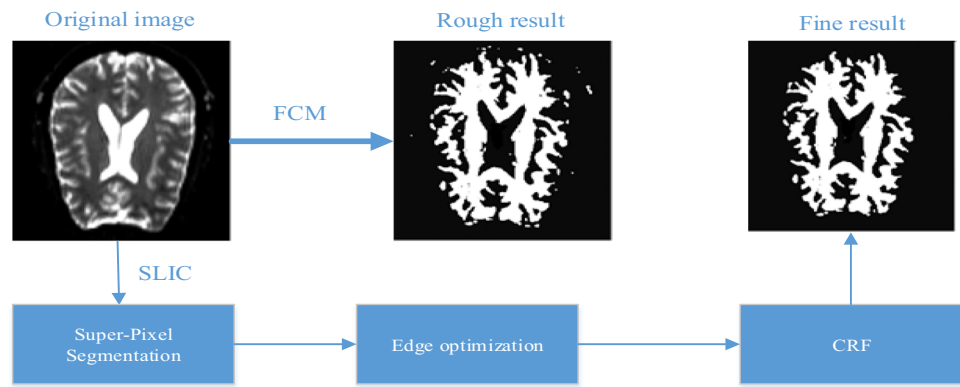
$$Q_i(x_i) = \frac{1}{Z_i} \exp \left\{ -\psi_u(x_i) - \sum_{i \neq j} [\psi_p(x_i, U_j)] \right\} \tag{2}$$

**Fig. 1** Brain white matter segmentation results for different receptive fields. **a** FCN-32 s; **b** FCN-16 s; **c** FCN-8 s; **d** benchmark





**Fig. 2** Framework of edge optimization algorithm



Formula (3) can be obtained by combining formula (2) and binary potential energy:

$$Q_i(x_i = 1) = \frac{1}{Z_i} \exp \left[ -\psi_u(x_i) - \sum_{l \in L} \mu(l, l') \sum_{m=1}^K \omega^{(m)} \sum_{i \neq j} k^{(m)}(f_i, f_j) Q_j(l') \right] \quad (3)$$

From Formula 3, we can get the specific solution of approximate probability distribution  $Q(X)$ . The flow chart of the mean field approximation algorithm for fully connected conditional random fields is summarized as follows:

- (1) Use formula (4) to initialize probability distribution Q:

$$Q_i(x_i) = \frac{1}{Z_i} \exp[-\psi_u(x_i)] \quad (4)$$

- (2) Determine whether the probability distribution Q converges or not. If it does not converge, repeat steps (a) to (d):
- (a) For all m, the information propagation values between all pairs of pixels are calculated:

$$\bar{Q}_i(l') = \sum_{i \neq j} k^{(m)}(f_i, f_j) Q_j(l') \quad (5)$$

- (b) Compatibility function is added to calculate its effect on probability distribution:

$$\hat{Q}_i(x_i) = \sum_{l \in L} \mu^{(m)}(x_i, l') \sum_{m=1}^K \omega^{(m)} \bar{Q}_i^{(m)}(l') \quad (6)$$

- (c) Locally renew formula (6) and solve the probability distribution of variable x:

$$Q_i(x_i) = \exp[-\psi_u(x_i) - \hat{Q}_i(x_i)] \quad (7)$$

- (d) Approximate probability distribution  $Q(X)$  can be obtained by normalizing  $Q_i(x_i)$

Steps (a) to (c) are the core steps of the mean field approximation algorithm. The complexity of steps (b) and (c) is linearly related to the number of variables in the probabilistic pattern model, which is efficient and less time-consuming. However, the relation between step (a) and the number of variables is a quadratic function because it is necessary to calculate the information propagation of other variables. In the field of image segmentation, the number of variables is generally the number of pixels in the image, so if this step is not optimized, the square complexity of the algorithm is unacceptable.

### Information propagation based on high-dimensional filtering

Approximate high-dimensional filtering can reduce the time complexity of information propagation process from square complexity to linear complexity. From the point of view of signal processing, the information propagation process can be represented by the convolution process of the Gauss kernel function  $G_A(m)$  in the feature space, as shown in Eq. (8):

$$\begin{aligned} \bar{Q}_i(l') &= [G_A(m) \otimes Q_j(l')](f_i) - Q_i(l') \\ &= \bar{Q}_i^{(m)}(l') Q_i(l') - Q_i(l') \end{aligned} \quad (8)$$

It is noted that the information propagation process of variables occurs only with variables except itself, so formula (8)

**Table 3** Configuration parameter

CPU	Intel(R) Core(TM) i3-6400 CPU 3.30GHz
RAM	32 GB
GPU	NVIDIA Titan Xp
Memory Size	12GB

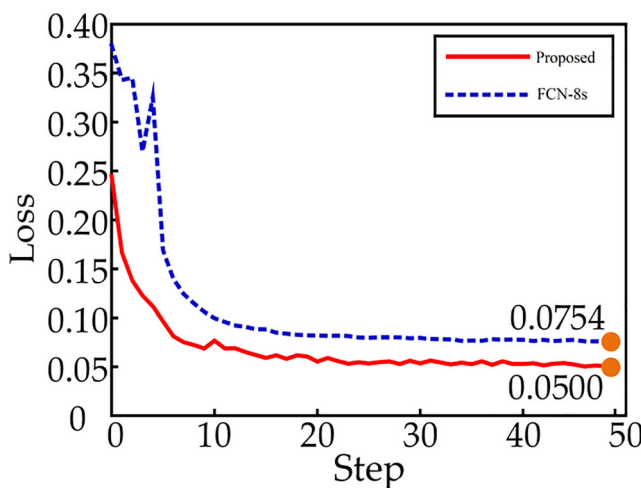
need subtract  $Q_i(l')$  on the basis of the stack of convolution functions  $\bar{Q}_i^{(m)}(l')$ . Essentially, the convolution process  $\bar{Q}_i^{(m)}(l')$  can be seen as a low-pass filtering process. When the distance is small, the convolution function is larger, resulting in larger weight. With the distance increasing, the convolution function becomes smaller, resulting in smaller weight. According to the sampling theorem, as long as the requirement of the number of sampling points is satisfied, the convolution function  $Q_i(l)$  can be reconstructed through a series of sampling points, in which the distance between sampling points is proportional to the standard deviation of the filter. In summary, the convolution process can be divided into three stages: firstly, acquire a part of the sampling points and get  $Q_i(l)$  by down-sampling  $Q(l)$ ; secondly, convolute sampling point  $\hat{f}$  :

$$\sum_{j \in \Phi} k^{(m)}(\hat{f}_i, \hat{f}_j) Q_j(l) \rightarrow \forall i \in \Phi \bar{Q}_i^{(m)}(l) \quad (9)$$

Finally,  $\bar{Q}_i^{(m)}(l)$  is obtained by up-sampling  $\bar{Q}_i^{(m)}(l)$ .

### Results and discussion

In order to verify the performance of the improved DTI white matter segmentation algorithm based on super-pixel and conditional random field proposed in this paper, we compare the



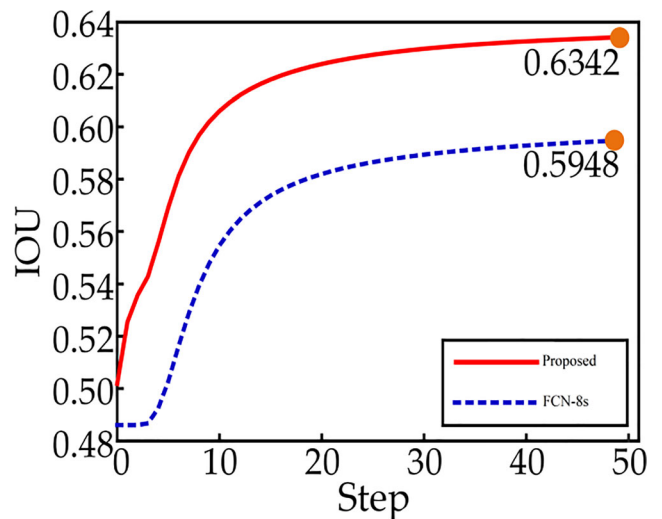
**Fig. 3** DTI data training process curve for LOSS

proposed algorithm with the current state-of-the-art segmentation algorithm, and analyze the segmentation results from both qualitative and quantitative perspectives, where we first introduce the preliminary preparation of the experiment, including the experimental platform, the data set, and the parameter setting strategy. Then the experimental verification of the improved algorithm is carried out, including the necessary verification of the algorithm steps and the comparison with other algorithms. Finally, the subjective and objective experimental data are given, and the necessary analysis is carried out.

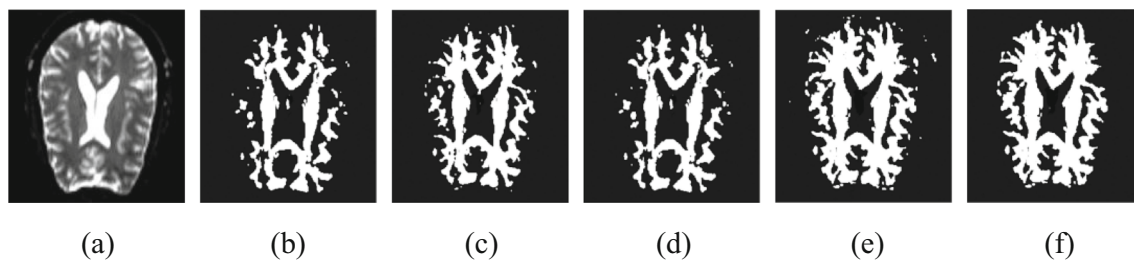
### Experiment setup

In order to ensure the final experimental data is convincing, all the algorithms are carried out under the same hardware and software conditions. The main configuration parameters of the computer used in this experiment are shown in Table 3:

NVIDIA Titan Xp is the most important part of the experiment in this article. The use of the GPU for the training and testing process of the FCN model greatly shortens the experimental time, and the SLIC algorithm implemented by the GPU can achieve the real-time processing effect when performing super pixel segmentation. In order to balance the software environment required for each experiment, the Ubuntu desktop 16.04 LTS operating system is adopted, where the deep learning framework Caffe is used to speed up the experiment. The self-built data-set and the public data-set are used for deep learning training and testing. The self-built data-set collects 7000 medical DTI image data from several hospitals and some data have been labeled by imaging specialists. The public data set is from the medical imaging data set DeepLesion released by NIHCC [28], which contains more than 32,000 lesion annotations from more than 10,000



**Fig. 4** DTI data training process curve for IOU



**Fig. 5** Brain white matter segmentation results for different model. **a** Original image; **b** FCN-8 s; **c** FCN-8 s + edge optimization; **d** FCN-8 s + CRF; **e** Proposed **f** Benchmark

cases. In order to facilitate training and testing, 12,150 images are selected as training samples and 8800 images as testing samples. The key parameters that need to be set in the improved algorithm mainly include the parameters of the FCN model, the number  $N$  of SLIC super-pixels and several hyperparameters  $w_1$ ,  $w_2$ ,  $\sigma_a$ ,  $\sigma_\beta$ ,  $\sigma_r$  in the fully connected conditional random field. Because the training of the FCN model in this paper uses the default parameters, the parameter settings of the FCN model are not discussed. We mainly discuss the number of SLIC super-pixels and the hyperparameters [29].

### Parameter optimization

This paper uses cross-validation to determine several hyperparameters for a fully connected conditional random field. First, we set two parameters  $w_2$  and  $\sigma_r$ . For these two parameters, they have little effect on the classification accuracy, more affect the smoothness, so the initial values are set as  $w_2 = 1$ ,  $\sigma_r = 1$ . According to the test results, the paper finally sets  $w_2 = 3$  and  $\sigma_r = 3$ . As for these three hyperparameters  $w_1$ ,  $\sigma_a$  and  $\sigma_\beta$ , we use an optimal search strategy from rough to fine. In addition, we select a small number of images to search on the training dataset. The initial values of these three parameters are set to  $w_1 = 3$ ,  $\sigma_a = 30$  and  $\sigma_\beta = 3$ . After the initial search range is set to a round of search, the search is researched within the range of the optimal value, and the incremental step is halved until the final search stops. This ensures that the conditional random field parameter set in this paper is the optimal parameter. After searching, the three values used in this article are  $w_1 = 5$ ,  $\sigma_a = 48$  and  $\sigma_\beta = 3$ . In addition, this paper sets the SLIC super-pixel to perform 10 iterations, and sets the CRF to perform 10 average field approximation iterations.

### Qualitative and quantitative analysis

In this paper, pixel accuracy (PA) and intersection over Union (IoU) are used to quantitatively analyze the segmentation accuracy. The meaning of PA is to calculate the ratio between the number of pixels accurately classified and the total number of pixels in the whole image, which is the most intuitive measurement index. The meaning of IoU is to measure whether an algorithm can effectively detect a specific category in an image. In addition to the above two evaluation indicators, this paper also uses the most widely used Mean Intersection over Union (MIoU) [30], which is the standard performance measurement index in the field of image segmentation. It is obtained by calculating the IoU of all categories.

The used metrics in this study is the IOU value of the segmentation [30, 31]. That is, the degree of overlap between the white-matter region of the segmented image and the white-matter region of the benchmark image, the loss function used in this study is binary cross entropy. The training process is iterated for 50 times. Figure 4 shows that the IOU values of DTI data and thoracic data in training set are respectively 0.6341 and 0.5947. The IOU values of these two different data are 0.6342 and 0.5948. As shown in Fig. 3, the Loss value of the verification set is 0.0754 and 0.05. By observing the image segmentation results, the DTI data segmentation had been well implemented (Fig. 4).

The core idea of the improved method in this paper is the application of super-pixel and the optimization of conditional random field model. As an improved algorithm, it is necessary to verify the improved algorithm to achieve the expected effect through the necessity analysis. Therefore, in order to verify the necessity of each step in the improved algorithm, we have designed some independent experiments to verify the effect and necessity of super-pixel edge optimization and CRF precise edge recovery. The normal FCN-8 s model is

**Table 4** Comparison of IoU for different models

Dataset	FCN-8 s	FCN-8 s + Edge optimization	FCN-8 s + CRF	Proposed
Self-built	62.7	65.9	69.6	75.2
DeepLesion	56.1	58.9	62.1	65.1



used as the benchmark for all comparison experiments; The super-pixel edge optimization is applied to the processing results of the FCN-8 s model as the first comparison experiment. The fully connected conditional random field is used as the post-processing step of the FCN-8 s model, but the super-pixel edge optimization is not used; The result of the FCN-8 s model is optimized through the super-pixel edge, and then the fully connected condition random field is used to further optimize the image edges. Figure 5 is an illustration of experimental results obtained on same test image.

From Fig. 5, we can see the comparison of the results for four independent experiments. First of all, from a subjective point of view, we can see that the super-pixel edge optimization is always better than the method without edge optimization. Figure 5b is closer to the real label than comparison algorithms. In addition, it can be concluded that the DTI segmentation precision of the edge can be improved by super-pixel edge optimization. For the FCN-8 s model combined with CRF, it has a significant improvement in subjective effect compared to the ordinary FCN-8 s.

We count the MIOU scores of the four experiments on two data sets, as shown in Table 2. The statistical data shows that the quantitative results are consistent with the conclusion of subjective analysis. The data shows that the edge optimization has increased the MIOU by 3.2% over the self-built dataset for the regular FCN-8 s model and by 2.8% over the DeepLesion dataset. Based on the above experimental results, it can be concluded that the DTI segmentation precision can be improved by super-pixel edge optimization. The purpose of using CRF is to reconstruct the edges more accurately. In the same situation, the necessity can be verified by comparing the performance of FCN-8 s with and without CRF as a post-processing step. As can be clearly seen from Table 4, CRF can always improve the MIOU score of the corresponding algorithm. Comparing Fig. 5d and e, the algorithm uses CRF to achieve accurate edge recovery, which is improved by 5% on the self-built data set and 4.1% on the DeepLesion data set. Edge optimization using super-pixels, and precise edge recovery using CRF are necessary, and both can improve the segmentation accuracy of the DTI white-matter segmentation algorithm.

## Conclusion

Diffusion tensor imaging (DTI) is a new imaging method that can be used to non-invasively measure the diffusion coefficient of water molecules in biological tissue structures in recent years. Since the DTI data is a tensor space, its segmentation is different from ordinary MRI images. Based on the existing deep learning model, an improved image semantic segmentation method based on super-pixels and conditional random fields is proposed. Firstly, this paper uses the existing

feature extraction model based on deep learning to obtain rough semantic segmentation results, including high-level semantic information of the image but lacking details of the image. In addition, the super-pixel segmentation algorithm is applied to obtain super-pixels that carries more low-level information. Secondly, due to the lack of image details in rough segmentation results, the segmentation of the edge of the image is inaccurate. In this paper, an boundary optimization algorithm is proposed to optimize the edge segmentation accuracy of the rough results. The edge segmentation effect in the rough results is preliminarily optimized using super-pixels. Finally, the use of super-pixels for local boundary optimization can improve the segmentation accuracy. Experiments results show that this segment is a practical and effective method.

**Acknowledgments** This work is financially supported by Liaoning Province Natural Science Fund(NO: 201602725) and Shenyang Science and Technology Bureau Fund (NO:1801331).

## Compliance with ethical standards

**Conflict of interest** We declare that we have no conflict of interest.

**Human and animal rights** The paper does not contain any studies with human participants or animals performed by any of the authors.

**Informed consent** Informed consent is obtained from all individual participants included in the study.

## References

1. Lawes, I. N. C., Barrick, T. R., Murugam, V., Spierings, N., Evans, D. R., Song, M., and Clark, C. A., Atlas-based segmentation of white matter tracts of the human brain using diffusion tensor tractography and comparison with classical dissection[J]. *NeuroImage* 39(1):62–79, 2008.
2. Leng, X., Fang, P., Lin, H., Qin, C., Tan, X., Liang, Y., Zhang, C., Wang, H., An, J., Wu, D., Liu, Q., and Qiu, S., Application of a machine learning method to whole brain white matter injury after radiotherapy for nasopharyngeal carcinoma[J]. *Cancer Imaging* 19(1):19, 2019.
3. Pengjiang Qian, Yizhang Jiang, Zhaohong Deng, Lingzhi Hu, Shouwei Sun, Shitong Wang, Raymond F. Muzic Jr., Cluster prototypes and fuzzy memberships jointly leveraged cross-domain maximum entropy clustering[J]. *IEEE Transactions on Cybernetics*, 2016, 46(1): 181-193.
4. Petronella, A., Vincken, K. L., Osch, M. J. P. V. et al., Probabilistic segmentation of white matter lesions in MR imaging.[J]. *Neuroimage* 21(3):1037–1044, 2004.
5. Honghao Gao, Wanqiu Huang, Yucong Duan, Xiaoxian Yang, Qiming Zou. Research on Cost-Driven Services Composition in an Uncertain Environment[J]. *Journal of Internet Technology(JIT)*, 2019,20(3):755-769.
6. Harris, G. J., Barta, P. E., Peng, L. W. et al., MR volume segmentation of gray matter and white matter using manual thresholding: Dependence on image brightness[J]. *AJNR Am. J. Neuroradiol.* 15(2):225–230, 1994.

7. O'Donnell, L. J., and Westin, C. F., Automatic tractography segmentation using a high-dimensional white matter atlas[J]. *IEEE Trans. Med. Imaging* 26(11):1562–1575, 2007.
8. Lao, Z., Shen, D., Liu, D., Jawad, A. F., Melhem, E. R., Launer, L. J., Bryan, R. N., and Davatzikos, C., Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine[J]. *Acad. Radiol.* 15(3):300–313, 2008.
9. Alfano, B., Brunetti, A., Larobina, M., Quarantelli, M., Tedeschi, E., Ciarniello, A., Covelli, E. M., and Salvatore, M., Automated segmentation and measurement of global white matter lesion volume in patients with multiple sclerosis[J]. *J. Mag. Reson. Imaging : JMRI* 12(6):799–807, 2000.
10. Valdés Hernández, M. d. C., Ferguson, K. J., Chappell, F. M. et al., New multispectral MRI data fusion technique for white matter lesion segmentation: Method and comparison with thresholding in FLAIR images[J]. *Eur. Radiol.* 20(7):1684–1691, 2010.
11. Gelineau-Morel, R., Tomassini, V., Jenkinson, M. et al., The effect of hypointense white matter lesions on automated gray matter segmentation in multiple sclerosis[J]. *Hum. Brain Mapp.* 23(21):4520–4530, 2012.
12. Ghafoorian, M., Karssemeijer, N., Heskes, T., van Uden, I. W. M., Sanchez, C. I., Litjens, G., de Leeuw, F. E., van Ginneken, B., Marchiori, E., and Platel, B., Location sensitive deep convolutional neural networks for segmentation of white matter hyperintensities[J]. *Sci. Rep.* 7(1):5110–5123, 2017.
13. Preetha Phillips, Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm[J], *Neurocomputing*, 2018, 272: 668-676
14. Hongwei, L., Gongfa, J., Ruixuan, W. et al., Fully convolutional network ensembles for white matter hyperintensities segmentation in MR images[J]. *NeuroImage* 53(8):119–123, 2018.
15. Yizhang Jiang, Kaifa Zhao, Kaijian Xia, Jing Xue, Leyuan Zhou, Yang Ding, Pengjiang Qian: A Novel Distributed Multitask Fuzzy Clustering Algorithm for Automatic MR Brain Image Segmentation[J]. *J. Medical Systems* 2019, 43(5): 118:1-118:9
16. Latini, F., Mårtensson, J., Larsson, E. M. et al., Segmentation of the inferior longitudinal fasciculus in the human brain: A white matter dissection and diffusion tensor tractography study[J]. *Brain Res.* 1675:102–115, 2017.
17. Labra, N., Guevara, P., Duclap, D., Houenou, J., Poupon, C., Mangin, J. F., and Figuerola, M., Fast automatic segmentation of white matter streamlines based on a multi-subject bundle atlas[J]. *Neuroinformatics* 15(1):71–86, 2017.
18. Sudre, C. H., Cardoso, M. J., and Ourselin, S., Longitudinal segmentation of age-related white matter hyperintensities[J]. *Med. Image Anal.* 38:50–64, 2017.
19. Qin, C., Guerrero, R., Bowles, C., Chen, L., Dickie, D. A., Valdes-Hernandez, M. C., Wardlaw, J., and Rueckert, D., A large margin algorithm for automated segmentation of white matter hyperintensity[J]. *Pattern Recogn.* 77:150–159, 2018.
20. Kubicki, M., Khan, U., Westin, C. F., Pieper, S., McCarley, R., and Shenton, M., Atlas-based white matter segmentation of DTI in schizophrenia[J]. *Schizophr. Res.* 102(1):82–82, 2008.
21. Bazin PL., Bogovic J., Reich D., Prince J.L., Pham D.L. Belief Propagation Based Segmentation of White Matter Tracts in DTI[C]. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009. Lecture Notes in Computer Science, vol 5761. Springer, Berlin, Heidelberg, 943-950
22. Kaijian Xia, Hongsheng Yin, Yu-Dong Zhang. Deep Semantic Segmentation of Kidney and Space-Occupying Lesion Area Based on SCNN and ResNet Models Combined with SIFT-Flow Algorithm[J]. *J. Medical Systems*, 2019, 43(1): 2:1-2:12
23. Klose, U., Erb, M., Saur, R. et al., Segmentation of the white matter in the brain based on MR-DTI fiber directions[J]. *Z. Med. Phys.* 15(4):247:259, 2005.
24. Vassar, R. L., Naama, B. G., Jessica, R. et al., Identification of neonatal white matter on DTI: Influence of more inclusive thresholds for atlas segmentation[J]. *PLoS ONE* 9(12):5426–5439, 2014.
25. Yu-Dong Zhang, Twelve-layer deep convolutional neural network with stochastic pooling for tea category classification on GPU platform, *Multimedia Tools and Applications*, 2018, 77(17): 22821-22839
26. Jonasson, L., Bresson, X., Hagmann, P., Cuisenaire, O., Meuli, R., and Thiran, J. P., White matter fiber tract segmentation in DT-MRI using geometric flows[J]. *Med. Image Anal.* 9(3):223–236, 2005.
27. Niogi, S. N., Mukherjee, P., and Mccandliss, B. D., Diffusion tensor imaging segmentation of white matter structures using a reproducible objective quantification scheme (ROQS).[J]. *Neuroimage* 35(1):166–174, 2007.
28. Bazin, P. L., Ye, C., Bogovic, J. A., Shiee, N., Reich, D. S., Prince, J. L., and Pham, D. L., Direct segmentation of the major white matter tracts in diffusion tensor images[J]. *Neuroimage* 58(2): 458–468, 2011.
29. Hao, X., Zygmunt, K., Whitaker, R. T., and Fletcher, P. T., Improved segmentation of white matter tracts with adaptive Riemannian metrics[J]. *Med. Image Anal.* 18(1):161–175, 2014.
30. Yizhang Jiang, Zhaohong Deng, Fu-Lai Chung, Guanjin Wang, Pengjiang Qian, Kup-Sze Choi, and Shitong Wang. Recognition of Epileptic EEG Signals Using a Novel Multi-View TSK Fuzzy System[J]. *IEEE Trans. Fuzzy Systems*, 2017, 25(1): 3-20.
31. Kaijian Xia, Hongsheng Yin, Pengjiang Qian, Yizhang Jiang, Shuihua Wang, Liver Semantic Segmentation Algorithm Based on Improved Deep Adversarial Networks in Combination of Weighted Loss Function on Abdominal CT Images[J]. *IEEE Access*, 2019 7: 96349-96358

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.