Journal of Medical Systems (2019) 43: 152
https://doi.org/10.1007/s10916-019-1289-2

IMAGE & SIGNAL PROCESSING

Brain Tumor Segmentation Based on Improved Convolutional Neural Network in Combination with Non-quantifiable Local Texture Feature

Wu Deng[1] · Qinke Shi[1] · Kai Luo[1] · Yi Yang[1] · Ning Ning[2]

Received: 3 February 2019 / Accepted: 11 April 2019 / Published online: 23 April 2019

Abstract

Accurate and reliable brain tumor segmentation is a critical component in cancer diagnosis. According to deep learning model, a novel brain tumor segmentation method is developed by integrating fully convolutional neural networks (FCNN) and dense micro-block difference feature (DMDF) into a unified framework so as to obtain segmentation results with appearance and spatial consistency. Firstly, we propose a local feature to describe the rotation invariant property of the texture. In order to deal with the change of rotation and scale in texture image, Fisher vector encoding method is used to analyze the texture feature, which can combine with the scale information without increasing the dimension of the local feature. The obtained local features have strong robustness to rotation and gray intensity variation. Then, the non-quantifiable local feature is fused to the FCNN to perform fine boundary segmentation. Since brain tumors occupy a small portion of the image, deconvolutional layers are designed with skip connections to obtain a high quality feature map. Compared with the traditional MRI brain tumor segmentation methods, the experimental results show that the segmentation accuracy and stability has been greatly improved. Average Dice index can be up to 90.98%. And the proposed method has very high real-time performance, where brain tumor image can segment within 1 s.

Keywords Tumor segmentation · Convolutional neural network · Non-quantifiable local feature;dense micro-block difference · Rotation invariant

Introduction

Malignant brain tumor is one of the most terrible cancers in the world. They often cause cognitive impairment. The most common brain tumors in adults are primary central nervous system lymphoma (PCNSL) and glioma, where glioma accounts for more than 80% of malignant tumors, so glioma is the focus of tumor segmentation [1]. However, since gliomas can appear anywhere in the brain and have irregular size and shape, it is still a challenging task to segment them [2]. Therefore, how to use modern pattern recognition technology to segment brain tumors efficiently and automatically has become an important research direction. Magnetic Resonance Imaging (MRI) technology is non-invasive tool and plays an important assistant role in brain tumor image segmentation. Clinically, MRI sequences of brain tumors include T1-weighted, T1C (Contrast enhanced T1 weighted images), T2-weighted images and FLAIR (Fluid Attenuated Inversion Recovery), which are usually combined with four kinds of images to diagnose the location and size of tumors [3, 4].

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

✉ Ning Ning
gkningning@126.com

Wu Deng
dengwu@wchscu.cn

Qinke Shi
shiqingke@wchscu.cn

Kai Luo
luokai@wchscu.cn

Yi Yang
yangyi@wchscu.cn

[1] Information Center, West China Hospital of Sichuan university, Chengdu 610000, Sichuan, China

[2] Department of Orthopaedics, West China Hospital of Sichuan University, Chengdu 610000, Sichuan, China

In recent years, researchers have used a variety of methods to segment brain tumor images, which can be generally divided into unsupervised learning algorithm and supervised learning algorithm. Most unsupervised learning algorithms belong to semi-automatic algorithms and the segmentation accuracy is generally low. Region growing method proposed by Kamnitsaset al. needs to set the initial points manually in the segmented image [5]. Szilagyi et al. proposed a cascade model based on multiple fuzzy C-means algorithms, but this algorithm was tested only on a limited data set, and the generalization of the model was not high [6]. Fuzzy C-means is used to select the initial contour of the tumor area, and then the level set method is used to extract the brain tumor boundary. Rakesh et.al proposed tumor texture segmentation using Dense Micro-block Difference (DMD), since the traditional DMD algorithm lacks the phase analysis of graphical element,so it could not better distinguish the same classified texture image formed by the rotation of graphical element. In addition, the incorrect selection of the initial contour will lead to unsatisfactory segmentationresults [4].

Compared with the unsupervised learning algorithm mentioned above, the overall segmentation accuracy of the supervised learning algorithm is improved because of the feedback mechanism of training sample. Compared with the traditional supervised machine learning algorithms such as Extremely Randomized Trees (ERT) and Support Vector Machine (SVM), the Convolutional Neural Network (CNN) [7] does not rely on manual feature extraction, but automatically learns high-level and complex features with task adaptive from training data, which makes the deep learning algorithm more accurate in tumor segmentation [8]. Havaei et al. proposed a cascaded CNN model to improve the segmentation accuracy of brain tumors, but the complexity of the network model is too high, which makes the training more difficult [9]. The segmentation accuracy can only reach about 85%. In order to speed up the training speed of the network in the CNN model, Pereira et al. adopted several convolution kernel with 3×3 small size to reduce the network parameters and deepen the number of layers so as to improve the accuracy of the model, but the segmentation accuracy is still not more than 90% [1]. Shi Dongli et al. combined with the fuzzy reasoning system, established learning rules to re-judge the probability of CNN predicting tumor points. Although the segmentation accuracy can reach 90%, but the algorithm has become semi-automatic [10]. Recently, it is difficult to improve the accuracy of brain tumor image segmentation based on CNN mainly because: (1) the labels of adjacent image patches are independent, without considering the correlation between labels [11]; (2) the size of the image patches limits the size of the receptive field, and the network can only extract local features, which makes CNN unable to segment fine tumor boundaries [12]; (3) CNN convolutes each adjacent image patches one by one, which has a high computational complexity. In view of the defect of CNN image segmentation, the Markov random field algorithm proposed in [13] and the conditional random field algorithm proposed in [14] and [15] take the correlation of image labels into account, and use it as a post-processing method of CNN for image boundary segmentation. Literature [16] proposed a full convolutional neural network that does not rely on the image patch to directly segment the pixel, which greatly reduces the computational complexity.

Combining the advantages of these algorithms, a fully automatic brain tumor MR image segmentation algorithm based on improved Fully Convolutional Neural Network (FCNN) andnon-quantifiable local feature is proposed. First of all, the concept of local feature with rotational symmetry is introduced. It is found that many rotation invariant local features are rotational symmetric to a certain degree. Therefore, we propose a novel local feature to describe the rotation invariant properties of the texture. In order to deal with the change of rotation and scale in texture image, Fisher vector encoding method is used to analyze the texture feature, which can combine with the scale information without increasing the dimension of the local feature. The resulting local features have strong robustness to rotation and gray intensity variation. Finally, the non-quantifiable local feature is fused to the FCNN to perform fine boundary segmentation to improve the segmentation accuracy of brain tumors.

Related works

Fully convolutional neural network

In 1998, LeCun et al. first applied CNN to the field of image recognition [11]. In 2012, AlexNet network structure proposed by Krizhevsky et al. made a great breakthrough in the field of CNN object classification [12]. In 2015, Long et al. proposed a full convolution neural network, a pixel-level image semantic segmentation network. This network is improved on AlexNet, and can classify each pixel end-to-end to achieve image segmentation task. The idea of extending theregionto arbitrary-sized inputs first appeared by Matan et al. [15], which extended the classic LeNet [11] to recognizenumbers and characters. Fully convolutional computation has also been exploited in the present era of many-layered networks, for example, sliding window detection by Sermanetet al [17], .semantic segmentation by Pinheiro and Collobert [18], and image restoration by Eigen et al. [5]. Fully convolutional training is rare, but used effectively by Tompsonet al [19]. to

learn an end-to-end detector and spatial model for pose estimation, although they do not discuss profoundly some new questions and situation [20–22].

FCNN converts AlexNet's two full-connection layers into convolution layers, so that the final output of the network is still a two-dimensional matrix, which retains the spatial information between pixels and is more conducive to feature extraction. In order to make the output and input sizes of the network consistent, the original FCNN directly sampled the output of the last convolution layer and restored it to the same size of the input image. However, this operation only utilizes the feature information of the fifth pooling, and theobtained features are rough, so it is difficult to achieve accurate segmentation of the target.

Our proposed algorithm will utilize the feature information of different pooling layers to realize multi-level fusion of feature information. After the output of the last convolution layer is up-sampled, the eigenvector matrix of the fourth pooling is fused for the first time, then the feature fusion matrix is up-sampled and the third pooling eigenvector matrix is fused for the second time, and then the fused eigenvector matrix is up-sampled and the second pooling eigenvector matrix is fused for the third time. Finally, the third fusion feature matrix is sampled up to get the same size feature matrix as the original image, so that a better feature information matrix can be obtained. In addition, the Batch Normalization (BN) layer is added after each convolution layer to accelerate the training speed of the network and improve the segmentation accuracy of the model. The network structure of multi-level feature information fusion is shown in Fig. 1. *Conv* is the convolution layer, which can extract edges, positions and other features of the image; *Pool* is the pooling layer, which can realize feature dimensionality reduction and retain the main features of convolution layer extraction; BN is the batch regularization layer, which can ensure that the weight distribution of the network does not change greatly after convolution; *Up* is the upper sampling layer, which mainly uses the deconvolution. *Crop_Fuse* cuts and fuses the feature matrix, *Pixelwise Prediction* is the prediction layer of pixel classification, and image segmentation is realized by classifying each pixel.

Dense micro-block difference feature

Dense Micro-block Difference (DMD) featureisdeveloped on the basis of key point descriptors such as BRIEF and ORB; however, unlike these key points, DMD features do not involve any threshold operations. The core idea of DMD feature is that every small image patch in texture image has a feature structure. In other words, if we can effectively obtain the structural features of sub-blocks, we can get a recognizable information. In order to encode the local structure of these sub-blocks, the gray difference of corresponding coordinates in the image blocks should be taken into account. In order to facilitate formal analysis, image patch p with size L × Land their corresponding coordinates can be expressed as $X = \{x_1, x_2, \ldots, x_K\}$ and $Y = \{y_1, y_2, \ldots, y_K\}$, so the local features of DMD are defined as

$$\upsilon(p, X, Y) = \{p(x_1) - p(y_1), \ldots, p(x_K) - P(y_K)\} \quad (1)$$

where $p(x)$ represents the pixel gray-level of the image sub-block in the coordinates $x = (a, b)^T$. In order to eliminate the influence of illumination changes, $l_2$norm regularization constraint is applied to the features of image blocks. It can be seen that the original formula of DMD is the average gray difference of image blocks. Unlike traditional DMD, we adopt a simplified single-scale formula and make great adjustment and optimization in the spatial selection of sampling coordinates. DMD feature selection adopts the sampling method mentioned in BRIEF descriptor [15]. The sampling points follow the isotropic Gauss distribution, namey, $(X, Y) \sim Gaussian(0, L^2/25)$. Because the distance and angle between sampling points are variable, the randomness of sampling coordinates helps to capture changes in different scales and directions.

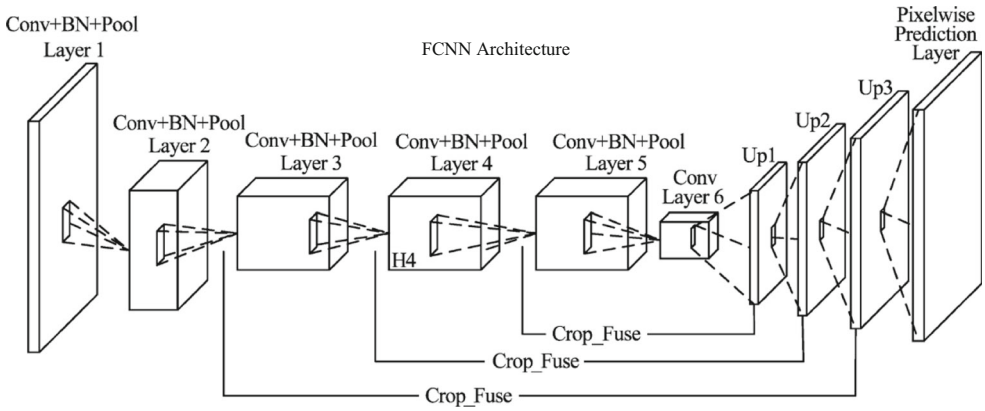

Fig. 1 The network structure of multi-level feature information fusion in FCNN architecture

Our proposed deep segmentation FCNN combined withnon-quantifiable texture feature

Improved FCNN model

In order to obtain good tumor characteristics, the CRF model is facilitated for fine boundary segmentation. In this paper, the original FCNN algorithm is improved, and a fine feature fusion model is proposed. In addition, in order to speed up the convergence of the network and improve the accuracy of the model, the BN layer is added to the network to form the IFCNN algorithm shown in Fig. 2. *Data* refersto the data layer, *Conv* is the convolution layer, *Pool* is the pooling layer, *ReLU* (Rectified Linear Units) is the activation function layer, *Up* is the upsampling layer, and *Prob* is the output probability layer. *Fuse* is a feature fusion operation for deep feature and multi-axis symmetric rotation invariant local feature. Under the premise of the same dimension and size of the high-dimensional feature matrix and the low-dimensional feature matrix, the data of the corresponding dimensions of the two target matrices are simply added. Crop indicates that if the dimensions and sizes of the two target matrices before the fusion are inconsistent (as shown in Table 1, the output size of the Pool_4 layer is $28 \times 28 \times 512$, and the output size of the Up_1 layer is $18 \times 18 \times 2$), the algorithm is to make the dimension and size of the two target matrices consistent. Perform a matrix cropping operation. When the dimension is inconsistent, the algorithm convolves the low-dimensional feature matrix with the convolution kernel with the size of $1 \times 1 \times N_{high}$, so that the two matrix dimensions are the same without changing the value of the low-dimensional feature matrix, where N_{high} is the number of high-dimensional feature matrices; When the size is inconsistent, the algorithm extracts the same size from the center of the low-dimensional feature matrix to the two sides symmetrically and reduces the loss of brain tumor feature informationbecause the feature information of the image is mainly concentrated in the middle part, so that the segmentation result is almost not effect.

Table 1 shows the specific parameters of the network of Fig. 2, where *Layer* is the name of each layer; Kernel indicates the dimension and size of the convolution kernel or pooled kernel; Stride indicates the step size of the sliding window during the convolution operation; Pad indicates the calculation before a certain edge expansion is done on the input matrix. In order to prevent the pooling feature information matrix from being too small, the original brain tumor image (240×240) is subjected to pad operation of 100; *Out-put Size* represents the dimension and size of the output feature matrix. The algorithm combines the feature matrices in different dimensions three times to make the acquired brain tumor features more fine. In addition, the BN layer can accelerate the convergence speed by keeping the weight distribution of the parameters in each iteration training unchanged.

Multi-axis symmetric rotation invariant local feature

Rotational symmetry structure is an important feature to achieve rotation invariance. The cyclic symmetry structures of LBP [6] and SRP [14, 23] make these local features become rotation invariants. Although it is not completely rotational symmetry, it has the rotational symmetry with N axis. N-axis rotation symmetry means that if the image is rotated around the center at an angle of 2 pi/N [24–26], the image structure will not change. However, coordinate pairs X and Y of DMD are randomly sampled from isotropic two-dimensional Gauss distribution, so there is no symmetry between sampling points. To introduce symmetry, coordinate points are sampled uniformly from a circular region, such as LBP and SRP. However, the use of circular structural coordinates loses randomness. Random is an important factor in DMD feature design. Researchers such as Calonder [15, 27, 28] have shown that the selection of random coordinates is more effective than the full symmetric circular distribution. In this paper, we strive

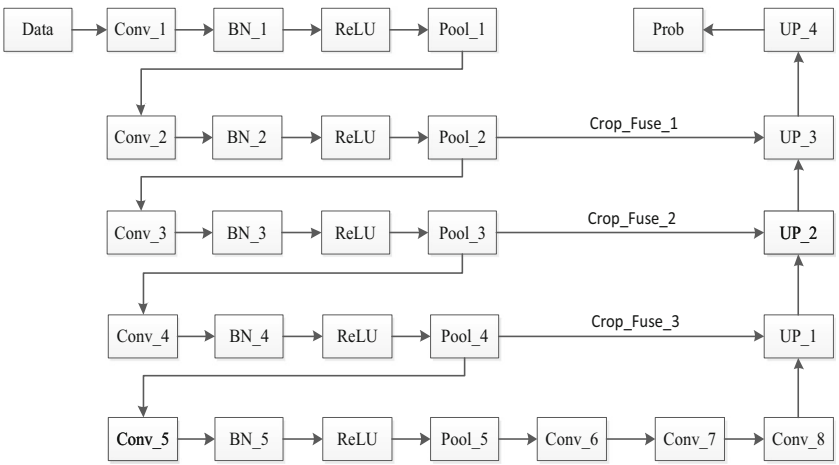

Fig. 2 Improved FCNN model

Table 1 The network structure parameters for our model

Layer	Kernel	Stride	Pad	Output Size	Layer	Kernel	Stride	Pad	Output Size
Data	–	–	100	440 × 440	Conv_6	7 × 7 × 4096	1	0	8 × 8 × 4096
Conv_BN_2	3 × 3 × 64	1	1	438 × 438 × 64	Conv_7	1 × 1 × 4096	1	0	8 × 8 × 4096
Pool_2	2 × 2 × 64	2	0	219 × 219 × 64	Conv_8	1 × 1 × 2	1	0	8 × 8 × 2
Conv_BN_1	3 × 3 × 128	1	1	219 × 219 × 128	Up_1	4 × 4 × 2	2	0	18 × 18 × 2
Pool_1	2 × 2 × 128	2	0	110 × 110 × 128	Crop_Fuse_1	–	–	–	18 × 18 × 2
Conv_BN_3	3 × 3 × 256	1	1	110 × 110 × 256	Up_2	4 × 4 × 2	2	0	38 × 38 × 2
Pool_3	2 × 2 × 256	2	0	55 × 55 × 256	Crop_Fuse_2	–	–	–	38 × 38 × 2
Conv_BN_4	3 × 3 × 512	1	1	55 × 55 × 512	Up_3	4 × 4 × 2	2	0	78 × 78 × 2
Pool_4	2 × 2 × 512	2	0	28 × 28 × 512	Crop_Fuse_3	–	–	–	78 × 78 × 2
Conv_BN_5	3 × 3 × 512	1	1	28 × 28 × 512	Up_4	16 × 16 × 2	4	0	324 × 324 × 2
Pool_5	2 × 2 × 512	2	0	14 × 14 × 512	Prob	–	–	–	240 × 240 × 2

to achieve a balance between randomness and symmetry of DMD coordinate pairs so as to achieve rotation invariance, while retaining the randomness of coordinates. Firstly, we propose a random coordinate method which can realize N-axis rotation symmetry. Then, using the symmetric structure of this coordinate, we propose a rotation invariant descriptor based on texture classification.

In order to achieve rotational symmetry in coordinate pairs, we first randomly select the initial coordinates from the isotropic Gauss distribution. The initial set contains only a fraction of the total number of coordinate pairs. These coordinates rotate at a fixed angle around the center, and a set of controllable coordinates can be obtained. Finally, combined with these controllable coordinate sets, the N-axis rotating symmetric coordinate sets are obtained. Through the above analysis, we can see that the proposed sampling point selection method not only satisfies the randomness of coordinate points, but also has structural symmetry. Next, we design a set of N-axis rotational symmetry with K-pairs coordinates in the form of mathematical model. Firstly, $T = K/N$ initial coordinate points are randomly selected from the isotropic Gauss distribution. These coordinate pairs are expressed as $X_0 = \{x_1^0, x_2^0, \ldots, x_T^0\}, Y = \{y_1^0, y_2^0, y_T^0\}$. The incremental rotation set of the initial random coordinates is shown in the following equation:

$$X_i = R_\theta \cdot X_{i-1} \tag{2}$$

where $i = (1, 2, \ldots, N-1)$, R_θ is a rotation matrix consistent with the angular change. The final set of rotating symmetric coordinates is a combination of a series of controllable coordinate sets, namely, $X_s = \{X_1, X_2, \ldots, X_{N-1}\}$. By using the same method and steps, the corresponding set of rotating symmetric coordinate $Y_s = \{Y_1, Y_2, \ldots, Y_{N-1}\}$ can be obtained. Figure 3 shows the comparison of random sampling of isotropic Gauss distribution in DMD algorithm with the results of multi-axis rotation sampling.

The main advantage of introducing a symmetry into the sampling coordinates is that the first rotation of the image patch can make the eigenvector shift linearly. Due to rotational

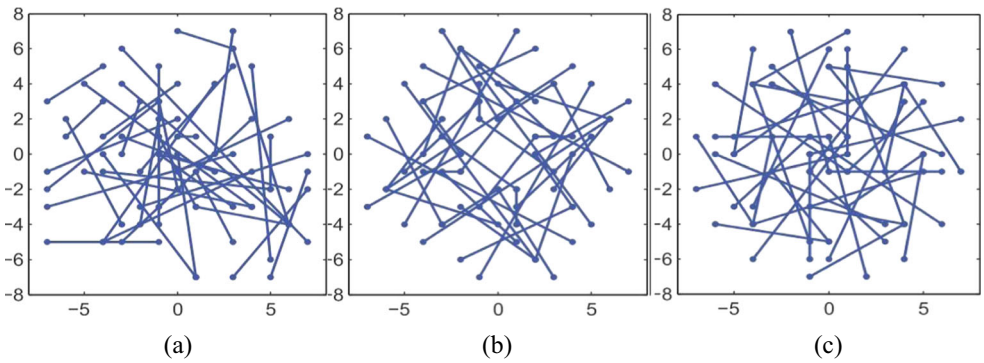

Fig. 3 Non-quantifiable local feature; (a) Random Sampling for Isotropic Gauss Distribution; (b)4-axis symmetric; (c)8-axis symmetric

change, linear shifts can be balanced by cyclic shifts of eigenvectors. In order to obtain the feature invariants under rotation, the eigenvectors need to be cyclically shifted so that the controllable subset corresponding to the max-sum feature will be in the first position. Thus, the subscript of the controllable set with the largest feature is written as follows:

$$\alpha = ar\,\mathrm{gmax} \sum_{j=1}^{T} \left(p\left(x_j^i\right) - p\left(y_j^i\right)\right) \tag{3}$$

it can be seen that the rotation invariant descriptor of image patch p can be calculated by cyclic shift of eigenvectors labeled as α subscripts. Since each controllable subset consists of T coordinates, the vectors are circularly shifted to positions. Then the cyclic shift of the eigenvector is expressed as

$$SDMD(p, Xs, Ys) - cyclic(\upsilon(p, Xs, Ys), \alpha * T) \tag{4}$$

$cyclic(A, P)$ represents a function that cyclically shifts the feature array A to its position P.

Improved fisher vector

Given a set of local features acquired from an image, they will be coded into vectors describing image features. In this paper, the improved Fisher vector [18] is used as the encoding method. The advantage of this method is that the features of low-dimensional space are mapped to high-dimensional space analysis through Fisher kernel, and the images of different scales have constant dimensions. The basic idea of Fisher vector is to construct a visual dictionary, which uses the logarithmic likelihood function gradient of model parameters to represent data, and then extracts features from the generated model. Fisher vectors use the Gauss mixture model (GMMs) to obtain a probability to represent local features. The first and second order difference between image descriptor and center of Gauss mixture model is obtained by coding. Compared with other coding methods, such as histogram and kernel codebook mapping, the learned higher-order statistics provide a more robust representation of texture images [29–31].

In order to obtain granularity at different resolutions, we calculate local features from image pyramids. As shown in literature [32], the local features of texture images will be extracted at scale 2^s, $s = -2 : 0.5 : 2$. The advantage of using multi-scale coding in coding stage is that the dimension of descriptor does not change with the change of scale s.

Experimental results and analysis

Data and evaluation criteria

In the experiment, BRATS 2015 (Brain tumor image segmentation benchmark) is used. Four modal MR images of FLAIR, TI, T2 and T1C of each patient in the database are all registered. In other words, the modal images are linearly aligned according to human standard brain, reflecting the same location of brain tumors. The three-dimensional size of each modal MR image is 240 × 240 × 155, and the true value label is the result of tumors manually calibrated by several experts. The experimental environment is as follows: CPU Intel Core i7 3.5GHz, GPU NVIDIA GeForce GTX1070, Ubuntu 16.04 LST 64-bit operating system, Caffe deep learning framework.

Dice Similarity Coefficient (DSC), Sensitivity and Positive Predictive Value (PPV) are used to evaluate the segmentation results of brain tumors. The similarity coefficient represents the similarity between the experimental segmentation result and the benchmark label, the sensitivity represents the proportion of the correct tumor points to the true tumor points after the experimental segmentation, and the positive predictive value represents the proportion of the correct tumor points to the tumor points after the experimental segmentation. Their expressionsare written as follows:

$$Dice = \frac{|P \cap T|}{(|P| + |T|)/2} \tag{5}$$

$$Sensitivity = \frac{|P \cap T|}{|T|} \tag{6}$$

$$Positive = \frac{|P \cap T|}{|P|} \tag{7}$$

where P is denoted as the segmentation result of this algorithm and T is denoted as the benchmark label of brain tumors segmented by experts.

Parameters setup and model training

The experiment randomly selected MR images of 100 patients with brain tumors as training sets. That is, the four modal voxel images with the size 240 × 240 × 155 are first sliced, and then the image preprocessing technique and data enhancement technique is used to obtain 51,000 Gy images with the size 240 × 240 asfusion training data. In this paper, the data enhancement processing is used to improve the accuracy of the training model.

In the training phase, the more mature Caffe deep learning framework is used to learn the model parameters of the training set. Considering that the Sigmoid and Tanh activation functions are easy to cause the gradient disappearance, the network selects ReLU as the activation function and uses the conventional Softmax classification loss function. Since the selection of learning rate is very important to the network, it determines whether the network can converge. The degree of convergence can neither be too big nor too small. This paper

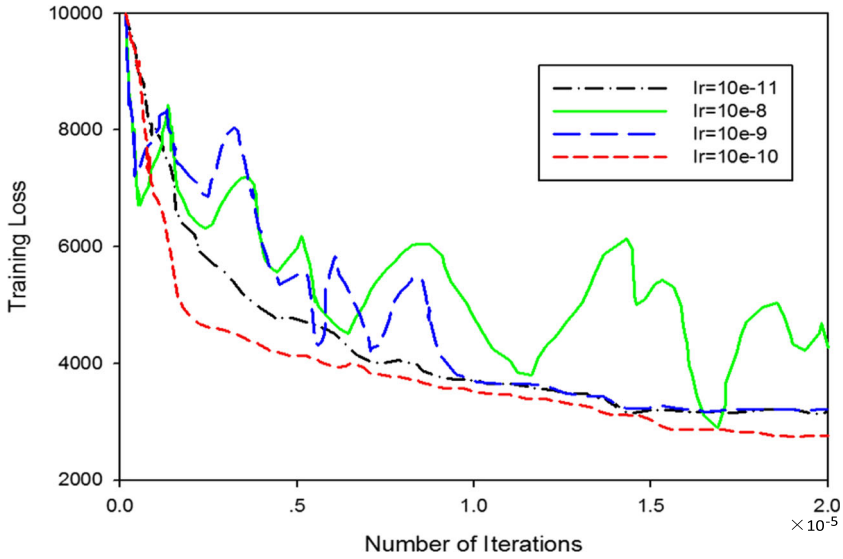

Fig. 4 The convergence degree of the network with different learning rates, where*lr* is learning rate

uses the original training set to compare the convergence of networks under different learning rates. Figure 4 reflects when the network selects a large learning rate such as $10e^{-08}$ and $10e^{-09}$, the loss of training cannot be converged. When the learning rate is $10e^{-10}$, the training loss can converge, and it has better convergence than that of $10e^{-11}$, so we select $10e^{-10}$ as learning rate.

The weight attenuation coefficient (WAC) has a certain influence on the degree of over-fitting of the network. This paper compares the training accuracy of the network under different weight attenuation coefficients. It can be seen from Table 2 that when the weight attenuation coefficient is too small, it does not play the role of weight attenuation. The accuracy of training is poor. When the weight attenuation coefficient is too large, the network has a certain over-fitting and the training accuracy is reduced. Therefore, the weight attenuation coefficient is 0.0005.

In order to verify the effectiveness of the BN layer, we use the original training set to compare the network training loss and training accuracy before and after adding the BN layer. The network added to the BN layer has been reduced to about 5000 after iteration of 8000 times, and the network has basically converged; while the network without BN layer has dropped to 5000 or so, it fully confirms that the BN layer can accelerate the convergence speed of the network, and the convergence loss value is small.

Table 2 Train accuracy of the network under different weight decay

WAC	Training accuracy	WAC	Training accuracy
0.0001	0.7348	0.0006	0.8428
0.0002	0.7654	0.0010	0.8125
0.0004	0.8401	0.0020	0.8012
0.0005	0.8562	0.0040	0.7568

Analysis of qualitative and quantitative experimental results

In order to verify the superiority of our improved algorithm, this paper compares it with the simple segmentation algorithm IFCNN which only fuses two features and the FCNN algorithm which does not fuse CRF. At the same time, the experiment also compares with these algorithmproposed by famous scholars Havaei and Pereia. Compared with the traditional CNN algorithm proposed by Havaei, it can be seen from Fig. 5 that Havaei algorithm has obvious over-segmentation, the boundary of brain tumors is not obvious and there are many isolated scattered points; Pereira algorithm has better accuracy and better performance than Havaei algorithm because of the deeper network layers, and reduces the over-segmentation of brain tumors to a certain extent. Smooth segmentation contour of brain tumors is obtained, but less feature fusion leads to less delicate boundary of brain tumors and low segmentation accuracy. Compared with IFCNN, FCNN has a certain enhancement in obtaining boundary information of brain tumors, but the segmentation boundary of brain tumors is still not delicate enough. Especially in the third case of complex brain tumors, because of the complexity of the boundary of the tumors, the segmentation results of various algorithms are unsatisfactory, but the algorithm in this paper can segment the more satisfactory results. Generally speaking, the FCNN model with more fine feature fusion and the end-to-end algorithm structure formed by the fusion of CRF can make similar pixels get the same label, so as to refine the boundaries of brain tumors and effectively solve the problem of over-segmentation and under-segmentation of brain tumors.

In addition, it can be seen from Table 3 that the proposed algorithm has higher segmentation accuracy than other algorithms. Compared with the conventional algorithms, our proposed algorithm improves 5.47% and 3.60% in Dice

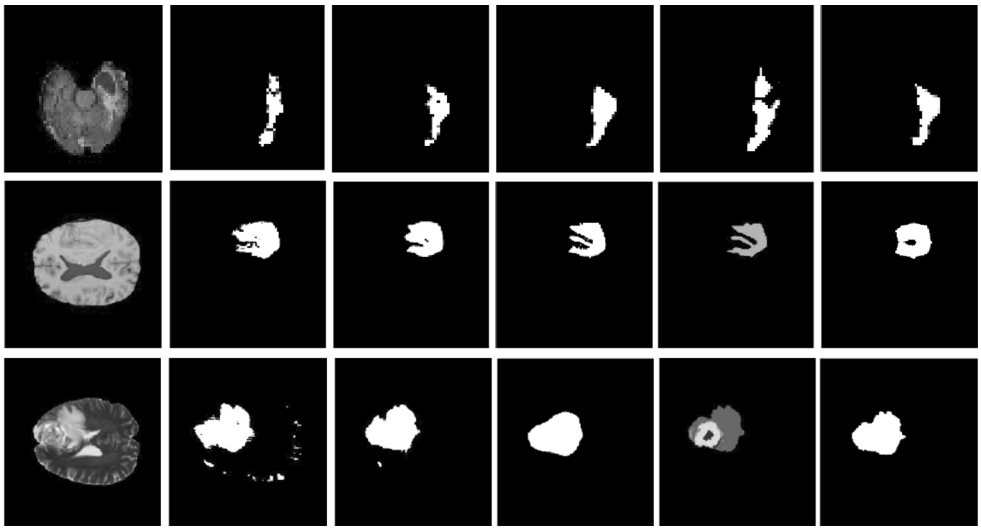

Fig. 5 Qualitative comparison for different algorithms; (**a**) original image; (**b**)FCNN; (**c**)IFCNN; (**d**)Pereiral; (**e**) Havaei; (**f**) Proposed

respectively, and the average improvement is 8.52% compared with the FCNN algorithm without fusion CRF. From the average time of segmentation for a brain tumor image, our algorithm has high real-time performance in the prediction stage, and can complete the segmentation of a brain tumor image in an average of 1 s.

In order to verify the stability of the algorithm, the experiment randomly selected 50 brain tumor slices from each patients, and evaluated each algorithm with Dice. Each image represents a patient's Dice value segmented by different algorithms, and each small rectangle box represents an algorithm. The height of rectangular box reflects the stability of the algorithm. The higher the rectangular box, the greater the fluctuation of Dice and the worse the stability of the algorithm. The horizontal line in the rectangular box is the average Dice value of the algorithm. It can be seen that the segmentation accuracy of the five algorithms is above 80%, but the performance of FCNN [33, 34] algorithm is poor both in segmentation accuracy and stability; Pereiral algorithm has slightly improved the segmentation accuracy and better stability than Havaei algorithm; IFCNN algorithm has greatly improved the segmentation performance compared with FCNN algorithm, and its stability is relatively poor compared with the comparison algorithm. However, our proposed algorithm, which combines multi-axis symmetric rotation invariant local feature and BN, not only improves the accuracy greatly, but also has a high stability with an average Dice of 91.29%. The overall performance of our proposed algorithm is better than other comparison algorithms.

Table 3 Segmentation performance evaluation for five different methods

Algorithm	Dice	Sensitivity	Positive	Time
Proposed	0.9129	0.9012	0.8916	1.0322
FCNN	0.8471	0.8533	0.8747	0.9867
IFCNN	0.8096	0.8326	0.8573	0.8356
Pereiral	0.8744	0.8964	0.8825	0.8935
Havaei	0.8525	0.8042	0.8412	2.7511

Conclusions

Accurate and reliable brain tumor segmentation is a critical component in cancer diagnosis. Build upon successful deep learning techniques, a improvedbrain tumor segmentation methods is developed by introduced fully convolutional neural networks and Dense Micro-block Difference Feature in a unified framework to obtain segmentation results with appearance and spatial consistency. First, we propose a novel local feature to describe the rotation invariant properties of the texture. In order to deal with the change of rotation and scale in texture image, Fisher vector encoding method is used to analyze the texture feature, which can combine with the scale information without increasing the dimension of the local feature. Theobtainedlocal features have strong robustness to rotation and gray intensity variation. Then, the non-quantifiable local feature is fused to the FCNN to perform fine boundary segmentation to improve the segmentation accuracy of brain tumors. Brain tumors occupy a small portion of the image, thus, deconvolutional layers are designed with skip connections to obtain a high quality feature map. Compared with the traditional MRI brain tumor segmentation methods of convolutional neural network, the experimental results show that the segmentation accuracy and stability has been greatly improved. Average Dice can be up to 90. 89%. And the real-time

performance of the proposed method is good. A brain tumor image can segmented with the training model within 1 s.In future, our proposed algorithm willintegrateinto the Intelligent Diagnosis System.

Compliance with Ethical Standards

We declare that we have no conflict of interest. The paper does not contain any studies with human participants or animals performed by any of the authors. Informed consent was obtained from all individual participants included in the study.

References

1. Pereira, S. et al., Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans. Med. Imaging:1–1, 2016.
2. Zhao, Xiaomei, et al. "Brain tumor segmentation using a fully convolutional neural network with conditional random fields." (2016).
3. Meier, R. et al., Clinical evaluation of a fully-automatic segmentation method for longitudinal brain tumor Volumetry. Sci. Rep. 6(1): 23376, 2016.
4. Rakesh, Z., and Kebin, J., Multiscale CNNs for brain tumor segmentation and diagnosis. Comput. Math. Methods Med. 2016:1–7, 2016.
5. Kamnitsas, K., et al. "DeepMedic for Brain Tumor Segmentation." 2016.
6. Szilagyi, Guotai, et al. "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks." (2017).
7. Ali, N., et al. Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods. Elsevier Science Publishers B. V.2016.
8. Dong, H., et al. "Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks." 2017.
9. Havaei, M., Davy, A., Wardefarley, D. et al., Brain tumor segmentation with deep neural networks.[J]. Med. Image Anal. 35:18–31, 2015.
10. Zhao, X. et al., A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. Med. Image Anal. 43:98–111, 2017.
11. Chang, P. D., Fully convolutional deep residual neural networks for brain tumor segmentation. In: International workshop on Brainlesion: Glioma, multiple sclerosis, stroke and traumatic brain injuries. Cham: Springer, 2016.
12. Sehgal, A., et al. "Automatic brain tumor segmentation and extraction in MR images." 2016 Conference on Advances in Signal Processing (CASP) IEEE, 2016.
13. Kaur, A.."An Automatic Brain Tumor Extraction System Using Different Segmentation Methods." Second International Conference on Computational Intelligence & Communication Technology IEEE, 2016.
14. Lang, L. Zhao, and K. Jia."Brain tumor image segmentation based on convolution neural network." International Congress on Image & Signal Processing IEEE, 2017.
15. Cabria, I., and Gondra, I., MRI segmentation fusion for brain tumor detection. Inform. Fus. 36:1–9, 2017.
16. Hussain, S., Anwar, S. M., Majid, M. "Brain tumor segmentation using cascaded deep convolutional neural network." Engineering in Medicine & Biology Society IEEE, 2017.
17. Qian, P., Zhou, J., Jiang, Y., Liang, F., Zhao, K., Wang, S., Su, K.-H., and Muzic, Jr., R. F., Multi-view maximum entropy clustering by jointly leveraging inter-view collaborations and intra-view-weighted attributes. IEEE Acc. 6:28594–28610, 2018.
18. Li, Y., Jia, F., and Qin, J., Brain tumor segmentation from multimodal magnetic resonance images via sparse representation. Artif. Intell. Med. 73:1–13, 2016.
19. Cimpoi, M., Maji, S., Kokkinos, I., Mohamed, S., and Vedaldi, A., Describing textures in the wild. Proc. IEEE Conf. Comput. Vis. Pattern Recog.:3606–3613, 2014.
20. Chang, P.D..."Fully Convolutional Deep Residual Neural Networks for Brain Tumor Segmentation."2016.
21. Calonder, M., Lepetit, V., Strecha, C., and Fua, P., Brief: Binary robust independent elementary features. Proc. Eur. Conf. Comput. Vis. (ECCV):778–792, 2010.
22. Rublee, E., Rabaud, V., Konolige, K., and Bradski, G., ORB: An efficient alternative to sift or surf. Proc. IEEE Int. Conf. Comput. Vis.:2564–2571, 2011.
23. Sharma, G., ul Hussain, S., and Jurie, F., Local higher-order statistics (LHS) for texture categorization and facial analysis. Proc. Eur. Conf. Comput. Vis.:1–12, 2012.
24. Sauwen, N. et al., Comparison of unsupervised classification methods for brain tumor segmentation using multi-parametric MRI. Neuroimage Clin. 12.C:753–764, 2016.
25. KJ Xia, JQ Wang, Y Wu. Robust Alzheimer Disease classification based on Feature Integration Fusion Model for Magnetic[J]. Journal of med. imaging and health infor. vol.7,1-6, 2017.
26. Chen, J. et al., WLD: A robust local image descriptor. IEEE Trans. Pattern Anal. Mach. Intell. 32(9):1705–1720, Sep, 2010.
27. Tang, Z., Wang, S., Huo, J., Guo, H., Zhao, H., & Mei, Y.. Bayesian Framework with Non-local and Low-rank Constraint for Image Reconstruction. Journal of Physics Conference Series. 2017.
28. Perronnin, F., Sánchez, J., and Mensink, T., Improving the fisher kernel for large-scale image classification. Proc. Eur. Conf. Comput. Vis.:143–156, 2010.
29. Varma, M., and Zisserman, A., A statistical approach to material classifi-cation using image patch exemplars. IEEE Trans. Pattern Anal. Mach. Intell. 31(11):2032–2047, Nov. 2009.
30. Quan, Y., Xu, Y., Sun, Y., and Luo, Y., Lacunarity analysis on image pat-terns for texture classification. Proc. IEEE Conf. Comput. Vis. Pattern Recog.:160–167, 2014.
31. Qian, P., Sun, S., Jiang, Y., Kuan-Hao, S., Ni, T., Wang, S., and Jr, R. F. M., Cross-domain, soft-partition clustering with diversity measure and knowledge reference. Pattern Recogn. 50:155–177, 2016.
32. Agn, M., et al. "Brain Tumor Segmentation Using a Generative Model with an RBM Prior on Tumor Shape. " 2016.
33. Qian, P., Chung, F., Wang, S., and Deng, Z., Fast graph-based relaxed clustering for large data sets using minimal enclosing ball. IEEE Trans. Syst. Man Cybern. B 42(3):672–687, 2012.
34. Hussain, S., and Triggs, B., Visual recognition using local quantized patterns. Proc. Eur. Conf. Comput. Vis.:716–729, 2012.