IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 25, 2020, accepted March 10, 2020, date of publication March 23, 2020, date of current version April 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2982197

Study on MRI Medical Image Segmentation Technology Based on CNN-CRF Model

NAIQIN FENG[1,2], XIUQIN GENG[3], AND LIJUAN QIN[1]

[1]School of Information Engineering, Zhengzhou University of Industrial Technology, Zhengzhou 451150, China
[2]College of Computer and Information Engineering, Henan Normal University, Xinxiang 453007, China
[3]Endocrine Department, Xinxiang Central Hospital, Xinxiang 453000, China

Corresponding author: Naiqin Feng (naiqinxinxiang@163.com)

This work was supported in part by the Key Scientific Research Projects of Higher Education Institutions of Henan Province under Grant 20A520039, in part by the Training Project of Young Backbone Teachers of Higher Education Institutions of Henan Province under Grant 2019GGJS279, and in part by the Science and Technology Research Project of Zhengzhou City under Grant 153PKJGG153.

ABSTRACT Image segmentation is an important technique for segmenting images without overlapping each other and having their own features. It has been rapidly developed in the field of medical imaging, but there is currently a difference between classification accuracy and segmentation accuracy for medical image segmentation. In this paper, the deep convolutional neural network is combined with the cascading structure, and a uniform learning framework is established with the use conditional random field. This paper first adds a cascading structure under the deep convolutional neural networks (DCNN) framework to more effectively simulate the direct dependencies between spatial closure tags. Secondly, the conditional random field (CRF) is used for post-segmentation processing, which effectively solves the contradiction between the segmentation accuracy and the network depth and the number of pooling times in the traditional convolutional network. Secondly, the CRF is used for post-segmentation processing, which effectively solves the contradiction between the segmentation accuracy and the network depth and the number of pooling times in the traditional convolutional network.

INDEX TERMS Image segmentation, deep convolutional neural network, cascade structure, conditional random field.

I. INTRODUCTION

In recent years, relevant research institutions have successively proposed algorithms such as normalized cut [1], graph cut [2], mean shift [3], and level set [4] to segment images, but the above algorithms are used for natural image segmentation. The effect is better, but it is not suitable for medical images with single gray value changes. It is limited when segmenting brain tumor MRI images [5]. Therefore, more and more researchers are studying new image processing techniques based on the segmentation of intracranial tumor images. Because deep convolutional neural networks (DCNN) have good adaptability to two-dimensional images, methods based on deep convolutional neural networks have become the main processing methods explored by researchers.

Le applied CNN, the convolutional neural network, to the first person in the academic field of image recognition. By directly inputting the original image to identify the visual laws, the pre-processing of complex images was eliminated. In recent years, this method has been widely used [6]; Havaei *et al.* used a cascade structure to provide the output of the first convolutional neural network model as an additional source of information for the next convolutional neural network model [7]; In order to improve the processing speed of the neural network, Pereira use a smaller convolution kernel in the convolutional neural network model, which not only improves the extraction speed, but also facilitates the acquisition of image features, but results in lower accuracy of segmentation. The impact cannot be ignored [8]; Zhao proposed a segmentation method based on multi-scale convolutional neural network model, which can extract the optimal size information of images and combine the pixel information of different scale regions, but the process is more verbose. The process is more complicated [9]; Li Xueqi and other scholars innovatively proposed a new auxiliary detection method, which uses the convolutional neural network algorithm to assist in the detection of small cell lung cancer, which combines the advantages of the previous detection algorithm and can also take into account the accuracy and Good to reduce the rate of misdiagnosis, improve learning efficiency,

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei.

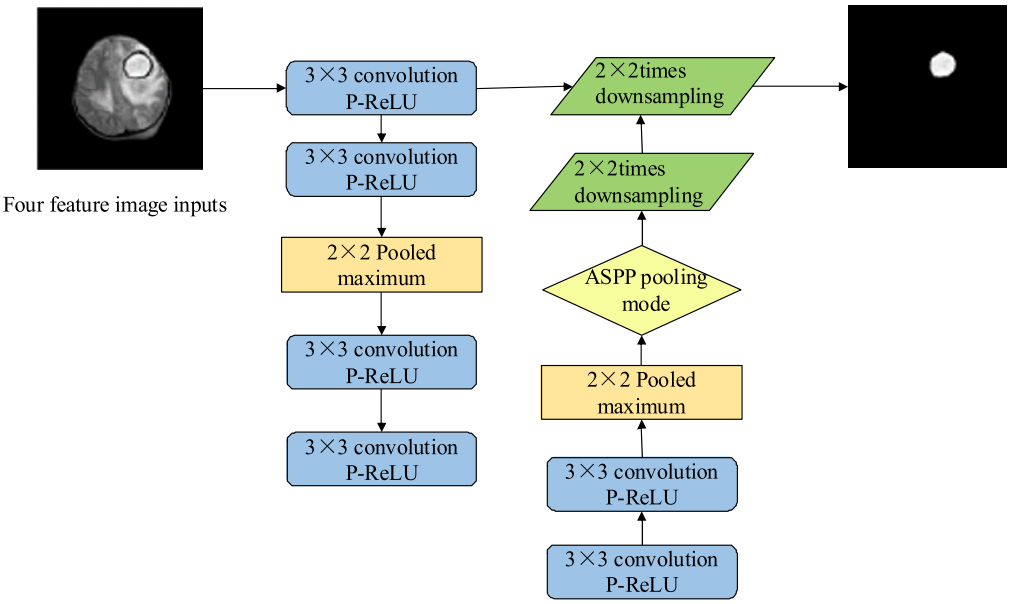


FIGURE 1. MRI segmentation method of brain tumor based on deep convolutional neural network.

can significantly improve the diagnosis rate, but the method needs to be better improved in terms of accuracy, there is still a great shortage of efficiency in the detection of 3D medical plain film [10]; Chen Hongxiang and others applied convolutional neural networks to the field of image semantic segmentation, and constructed and implemented a deep neural network structure combining convolutional neural networks and deconvolutional neural networks to directly predict images at the pixel level. The semantic category to which it belongs, but the scope of the model will be limited, resulting in lower accuracy of complex image edge regions [11]; Wang Yuanyuan *et al.* used convolutional neural networks for computer-assisted diagnosis of lung tumor positron emission tomography (CT) computed tomography, which not only provides accurate quantitative analysis to compensate for human eye inertia and gray-scale insensitive defects. It can also assist doctors in accurate diagnosis and treatment, but it may cause gradient dispersion due to the network too deep, resulting in reduced segmentation accuracy [12]. Li Fang and others studied the application of convolutional neural networks in brain tumor image segmentation. Using the characteristics of the cascade of multiple convolutional layers under the same receptive field, the nonlinearity is much larger than that of a single convolutional layer. The cascade connection layer joins the network and adds identity mapping to promote gradient flow and improve training speed. The convolutional layer constitutes a full convolutional neural network to achieve higher-level brain tumor segmentation. Studies have proved that the model can obtain better segmentation results, but the method is complicated to operate and increases the doctor's work time [13].

In this paper, the method of convolutional neural network combined with cascading structure can effectively and accurately realize the automatic segmentation of MRI images of brain tumors, saving doctors' working time and improving diagnosis efficiency. The first chapter introduces in detail the hierarchical structure of deep convolutional neural network and the image segmentation processing framework with residual structure with residual structure. The second chapter introduces the method and necessity of coupling the conditional random field (CRF) model with the deep convolutional neural network; In the third chapter, based on the proposed new segmentation method of MRI segmentation of brain tumors, combined with feature extraction and convolutional neural network and applied to the diagnosis of brain tumors, the automatic segmentation of MRI images is realized, and the segmentation of CNN-CRF is proposed. Method structure diagram; the fourth chapter carries out simulation experiments for different image segmentation, and the experimental results are compared. The use of deep convolutional neural network algorithm to process medical images is a development trend in the medical field, and it can provide better theoretical support for future image segmentation techniques.

II. FRAME OF IMAGE SEGMENTATION PROCESSING BASED ON DCNN ALGORITHM

A. DEEP CONVOLUTIONAL NEURAL NETWORK (DCNN) FRAMEWORK

The neural network of this segmentation algorithm adopts the sub-network structure shown in Figure 1 below. In order to improve the efficiency of multiple convolutional networks, it is also possible to more directly simulate the dependence of adjacent tags in the segmentation, so that the final prediction energy can be Influenced by the model's value on nearby tags, this DCNN uses a cascade structure to implement the output of the first convolutional network as an additional input to the second convolutional network. At the same time, in order to enable DCNN to train more effectively, a residual connection can be used, that is, an independent connection bypassing the parameter layer is created in the network [14].

1) COLOR/GRAYSCALE FIGURES

The convolutional layer involves three tensors: two feature maps and one convolution kernel, as in equation (1):

$$x'_j = \sum_{i \in M} x^{j-1} \otimes k'_{ij} + b'_j \tag{1}$$

where x^{j-1} is the input feature map, k is the convolution kernel, b is the deviation term, and x_j is the output feature map of the convolution operation.

The convolution process can be understood as the convolution kernel covering the original image or covering on the upper layer, and outputting the local region of the feature map, and weighting and summing each position of the covered region [15].

2) POOLING LAYER

The pooling layer is an important part of the convolutional neural network, which can reduce the complexity of the operation and reduce the over-fitting problem of high-dimensional features by reducing the connection between the convolutional layers [16]. The commonly used pooling operation is the maximum pooling operation:

$$x'_j = \max(x^{j-1}) \tag{2}$$

The pooled value of this area is the maximum value that passes through the calculation area. The maximum pooling can reduce the offset of the estimated mean due to convolutional layer parameter errors, thereby retaining more of the texture information of the image. Since the random pooled gradient back-transmission is more complicated, and the network often needs to consider the local texture details, the maximum pooling algorithm is often used.

3) ACTIVATION LAYER

In order for CNN to fit a nonlinear map, an activation function can be used to incorporate nonlinear changes. Sigmoid or Tanh activation functions are commonly used nonlinear activation functions of neural networks, and often produce robust optimization results in the training of deep neural networks [17]. The Sigmoid function is:

$$S(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

The Tanh activation function converges faster, and the function is:

$$Tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \tag{4}$$

The images of the above two activation functions are shown in Figure 2:

Both of these functions have gradient problems due to supersaturation.

4) DECONVOLUTION LAYER

In order to make the final result maintain the size of the input image, the input needs to be de-pooled and then convoluted [18]. The deconvolution operation is shown in Figure 3.

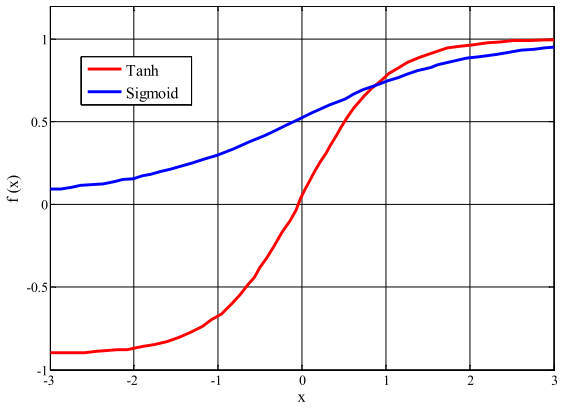


FIGURE 2. Sigmoid and Tanh activation functions.

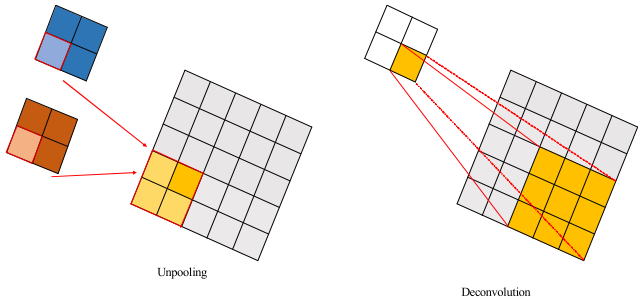


FIGURE 3. Deconvolution operation.

The deconvolution process can be summarized as:The position information of the activation value is tracked through the unpooling operation, and the input is enlarged once to obtain a sparse picture; Deconvolution is similar to bilinear interpolation, and can selectively ''fill'' unpooling the enlarged sparse results.

B. LOSS LAYER

In order to reverse the gradient and adjust the network weight, it is also necessary to give the loss function between the supervisory information and the network predictive output. Scale-invariant loss function [19], [20] as in equation (5):

$$L(y, y*) = \frac{1}{n} \sum_i d_i^2 - \frac{\lambda}{n^2} (\sum_i d_i)^2 \tag{5}$$

The loss function considers both the Euclidean distance loss and the depth network to preserve the 3D relative structure of the scene as much as possible when predicting the depth, which reduces the sensitivity of the CNN estimation error to image scene changes to some extent.

C. CASCADE STRUCTURE

The traditional convolutional neural network has a disadvantage, that is, its prediction for each segmentation label is separate, so it cannot effectively simulate the direct dependency relationship between the space closing labels. Dependence, so that the final prediction can be affected by the model's value on nearby labels. This paper uses a cascade structure [21], [22]. The cascade structure determines the number and depth of neurons by itself, and the original structure can be maintained after the training set changes. The advantages are also obvious. The cascade can organically combine

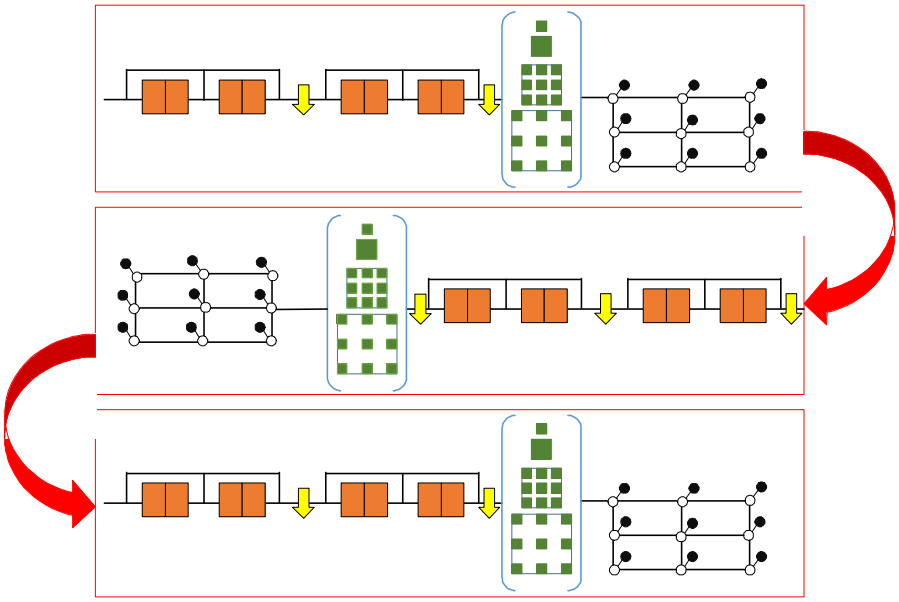


FIGURE 4. Schematic diagram of the cascade structure.

various methods to complement each other and improve the overall system. performance.

The cascading structure used in this paper enhances the division of labor for each convolution, and solves different problems step by step [23]. The first CNN is only responsible for segmenting the part corresponding to the tumor feature from the original image, and using the output segmentation result as the input data of the second CNN, and the second (and subsequent) CNN will be further segmented. The tumor continues to be segmented into specific substructures. In the test data set of this paper, the sub-tumor structure is divided into four parts: necrosis, edema, enhancement of tumor nucleus and unenhanced tumor nucleus. In order to simplify subsequent segmentation tasks, a small area that needs to be segmented can be cropped from the global. In training, this area is generated from standard data, and in the test, this area is generated from the results of the previous segmentation task. The specific structure is shown in Figure 4:

D. RESIDUAL CONNECTION

In order to make the training process of the deep convolutional neural network more efficient, this segmentation method uses a residual network, that is, creates an independent connection in the network that bypasses the parameter layer. The residual network is easier to optimize than the ordinary network, and can improve the network performance by simply increasing the network depth, and at the same time solve the degradation problem caused by the increased depth. The network in this paper contains three residual blocks, each with two convolutional layers and a residual connection bypassing them. The input to the residual block is added directly to the output to encourage the block to learn the residual function based on the input. This method can make the information spread more smoothly, accelerate the training

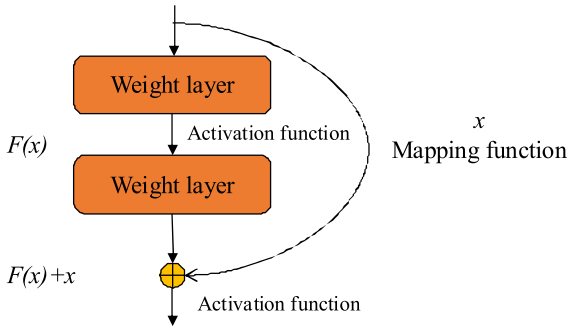


FIGURE 5. Schematic diagram of the residual network.

convergence, and avoid the gradient dispersion hinder the training deepening [24].

The specific structure of the residual block is shown in Figure 5.

It can be seen that x is the input value of this layer of residual block, and $F(x)$ is the output that is linearly changed and activated after the first layer of weights, also called $F(x)$ as the residual. As shown in the residual network, after the second layer undergoes a linear change, $F(x)$ adds this layer of input value X, and then activates and outputs.

E. ATROUS SPATIAL PYRAMID POOLING STRUCTUR(ASPP)

Void convolution (aka expanded convolution) was originally used to improve the computational efficiency of unsampled wavelet transforms and was used in DCNN for Sermanet. In the DCNN whose structure has been determined, it is necessary to input a fixed-size picture. When detecting pictures of various sizes, it will undergo a series of operations such as cropping or zooming, but this often causes a decrease in the accuracy of recognition detection, so this is targeted at this point. Holschneider proposed the ‘‘atrous space pyramid pooling’’ method. The advantage of this algorithm is that it can enable the network we construct to input images of any

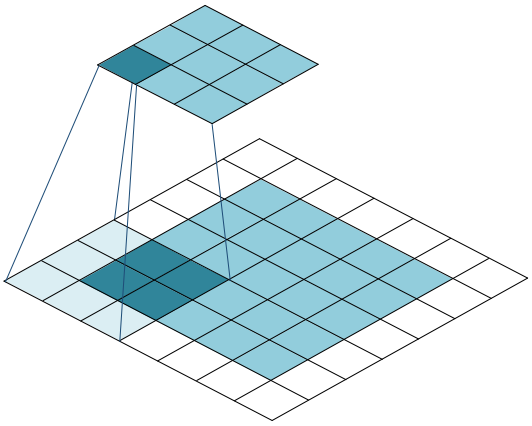


FIGURE 6. Example of cavity convolution.

size without the need for cropping and scaling, and at the same time effectively improve image accuracy [25].

Although the cavity convolution increases the filter, it does not increase the parameters and the amount of computation, so the spatial resolution of the CNN feature map response can be controlled simply and accurately. At the same time, the cavity convolution can expand the convolution kernel without increasing the number and calculation of non-zero parameters, thus effectively controlling the size of the receiving field, thereby achieving accurate segmentation (small scale) and region classification (large scale). Balance between. Figure 6 shows a hole convolution with a convolution rate of 2, which samples only 3 × 3 nodes in the range of 5 × 5 nodes.

Inspired by the spatial pooling pyramid method of circular CNN, Chen proposed an atrous space pyramid pooling model (ASPP) model based on cavity convolution. The model uses hollow convolutional layers with different parameters to sample and extract the features extracted at each sampling rate. Finally, the results are fused to obtain feature information of different scales, and the targets are accurately and effectively classified. At the same time, in order to avoid too large parameters will lead to a small proportion of effective filtering problems, Chen proposed an improved method based on the original ASPP. In brief, it is in the depth of a convolution, using the global average pooling method, using 1 × 1 convolution and 3 × 3 hole convolution of three different parameters (the batch is normalized after each convolution) The method of combination. As shown in Figure 7.

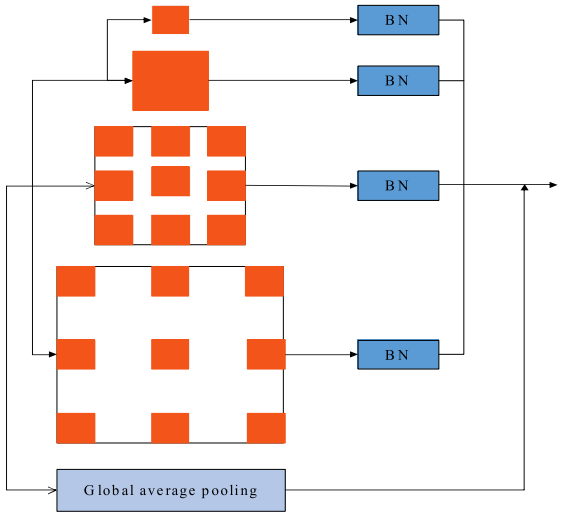


FIGURE 7. ASPP structure example.

III. CONDITITIONAL TANDOM FIELD (CRF) MODEL

Usually the graphs generated by CNN for regression are relatively ambiguous, and CRF can improve this shortcoming by means of conditional probability modeling.

CRF is a conditional random field, which is a conditional probability distribution model of another set of output random variables given a set of input random variables. It is a discriminative probability undirected graph model, and the discriminant is the conditional probability distribution. mold. CRF is mostly used in the field of natural language processing and image processing. In NLP, it is a probabilistic model for labeling and dividing sequence data. According to the definition of CRF, the relative sequence is the given observation sequence X and the output sequence Y, and then passed. Define the conditional probability $P(Y|X)$ to describe the model. In a linear chain conditional random field, in the case of a given observation sequence X, the probability of a particular sequence Y is $P(Y|X)$, which by definition is:

$$P(Y|X) = \exp\left(\sum_{i,k} \lambda_k t_k(Y_{i-1}, Y_i, X, i) + \sum_{i,l} \mu_l s_l(Y_i, X, i)\right) \quad (6)$$

where $t_k(Y_{i-1}, Y_i, X, i)$ represents the transfer function, $s_l(Y_i, X, i)$ represents the state function, and λ_k and μ_l are the weights of the two functions, respectively.

The output random variable of CRF is assumed to be an undirected graph model or a Markov random field, and the input random variable is not assumed to be a Markov random field. The graph model structure of CRF can theoretically be given arbitrarily, but we are common. It is a special conditional random field defined on a linear chain, called a linear chain conditional random field.

In DCNN, there is a contradiction between classification accuracy and segmentation positioning accuracy: in order to better complete the classification task, a deeper network and more pooling are needed, but this will lead to an increase in uncertainty and deep node receptive fields. In fact, it is not conducive to accurately locate and smooth the results, so as to correctly segment the boundary, coupling the conditional random field to the DCNN can alleviate this contradiction to some extent [26].

The CRF is a model evolved from the Markov Random Field (MRF), which is an undirected graph in which each node satisfies the Markov property [27]. If an MRF has only two variables, one of which is the output of another variable under a given condition, then this MRF is called a CRF. Unlike short-range CRFs used to smooth out noisy segmentation maps in previous weak classifiers, modern deep convolutional neural networks typically do not require smoothing, smoothing and may even impede the accuracy

of the results. Therefore, the contrast-sensitive local CRF potential function is no longer used, but a fully connected potential function is used.

$$E(X) = \sum\nolimits_i \theta_i (x_i) + \sum\nolimits_{ij} \theta_{ij} (x_i, x_j) \tag{7}$$

where x is the label assigned by the pixel, and its one-dimensional potential energy $\theta_i (x_i) = -\log P (x_i)$ is the probability of the distribution of the pixel i calculated by the DCNN on the label. Paired potential energy has a valid form of inference using a fully connected graph:

$$\begin{aligned}\theta_{ij} (x_i, x_j) &= \mu (x_i, x_j) \left[w_1 \exp \left(-\frac{||p_i - p_j||^2}{2\sigma_\alpha^2} - \frac{||I_i - I_j||^2}{2\sigma_\beta^2} \right) \right. \\ &\left. + w_2 \exp \left(-\frac{||p_i - p_j||^2}{2\sigma_\gamma^2} \right) \right]\end{aligned} \tag{8}$$

where $x_i \neq x_j$ is $\mu (x_i, x_j) = 1$, otherwise 0, that is, only the connection points of two different tags will be punished. Other expressions use two Gaussian kernels in different feature spaces. The first "bilateral" core depends on the position of two pixels (represented by p) and the intensity of each scan (indicated by I), while the second core depends only on the position of the pixel. The first core marks pixels of similar intensity and position with similar labels, while the second core only needs to consider spatial proximity in smoothing. The calculation method of the potential function is also very similar, and it is more convenient and effective in practical applications.

IV. SEGMENTATION METHOD BASED ON CNN-CRF

Based on the novel segmentation method of MRI segmentation of brain tumors proposed in this paper, combined with feature extraction and convolutional neural network and applied to the diagnosis of brain tumors, the automatic segmentation of MRI images is realized, aiming at the size, shape and position of brain tumors. Differently, the filtered images with different modal features are segmented by CNN, and the segmented results are fused. The fused results are post-processed based on GMM model, and finally the segmentation results are generated [28], [29].

This method is more accurate and better than the conventional method, but at the same time it uses the basic convolutional neural network, that is, CNN has obvious deficiencies, the network structure is backward and can not effectively simulate the direct dependence between the spatial closure tags. relationship. In response to this, the segmentation of the convolutional neural network is modified, and the image segmentation is performed using a DCNN with cascade structure, including three residual layers and a ASPP model. The pooling method, in order to better solve the contradiction between the segmentation accuracy and the network depth and the number of pooling times in the traditional deep convolutional network, the CRF is used to perform the post-processing of the segmentation result.

TABLE 1. Algorithm performance comparison table.

Algorithm	Dice	Positive predictive value	Sensitivity
Fuzzy clustering	0.82	0.8572	0.8103
FCM clustering	0.76	0.7822	0.8762
SVM	0.85	0.8253	0.8274
CNN	0.78	0.8427	0.8521
CNN-CRF	0.89	0.8702	0.8784

In order to better solve the contradiction between the segmentation accuracy and the network depth and the number of pooling times in the traditional deep convolutional network, the CRF is used to segment the result. The following figure is the modified deep convolution network. The brain tumor segmentation method is upgraded in the convolutional neural network segmentation compared to the original method structure, and the conditional random field model is used in the post-processing part to avoid mis-segmentation of the previous fusion results.

The optimization process optimizes the weights and deviations between the various layers of the CNN to minimize the error function, which is to make the CNN prediction results most consistent with the actual situation.

In this paper, the adaptive instantaneous estimation algorithm (Adam algorithm) is used to optimize the parameters. The Adam algorithm is an improved algorithm based on the Adadata algorithm. The Adadata algorithm improves the problem of the general learning rate of the algorithm decreasing. This method does not need to accumulate the square gradient, but limits the window of the previous gradient to w, and calculates different learning rates for each parameter, effectively preventing the learning rate attenuation or gradient disappearing. On this basis, the Adam algorithm sets different parameters for each momentum. It not only stores the exponential decay average of the previous square gradient, but also maintains the exponential decay mean M(t) of the previous gradient, namely:

$$\theta_{t+1} = \theta_t - \frac{\eta \cdot \hat{m}_t}{\sqrt{\hat{v}_t + \varepsilon}} \tag{9}$$

where A and B respectively represent the first-order matrix deviation (momentum term) and the second-order matrix deviation:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{10}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{11}$$

where β_1 and β_2 are the exponential decay rates of the matrix, respectively. In practical applications, the Adam algorithm has faster convergence speed and higher learning efficiency than other adaptive learning rate algorithms. At the same time, some problems of other optimization methods are also avoided, such as the disappearance of the learning rate, the slow convergence rate or the loss of the function function caused by the parameter update of the high variance.

At the same time, in order to highlight the advantages of this algorithm, several evaluation algorithms are used to compare and analyze the image segmentation. It can be cut from Table 1. The Dice coefficient, accurate prediction value

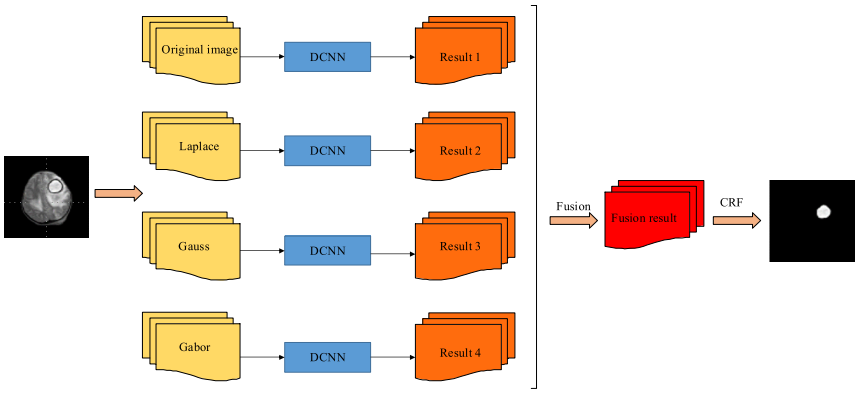


FIGURE 8. CNN-CRF segmentation method structure diagram.

and sensitivity of this algorithm are higher than other algorithms. Therefore, it proves that this algorithm has certain advantages.

V. SIMULATION

A. PARAMETER SETTINGS

In this paper, the learning rate is set to 10^{-3}, the block size is 5, and the maximum number of iterations is 2×10^4,based on a large number of training sets and experience from experiments, using advanced equipment in the laboratory as support, so it is worth a good foundation to solve these problems. The brain MRI image in this article is taken from the data in the hospital CT,a total of three neural network systems were trained in the experiment. They were a DCNN-based target method, a control network 1 without an ASPP structure, and a control network 2 without a CRF. In this experimental stage, four segmentation methods were used to conduct comparative experiments to prove the improvement of the quality of the two segmentation techniques [30].

B. SEGMENTATION RESULT

Figure 9 below shows the segmentation results of the segmented tumor MRI images, where (a) represents the original brain MRI image, and (b) represents the results of expert manual segmentation, (c), (d), (e), (f) Respectively represent the segmentation result of the target method, the segmentation result of the target DCNN, the segmentation result of the comparison network 1, and the segmentation result of the comparison network 2 (where the control network 1 is not using the ASPP structure, and the comparison network 2 is not adopted CRF method).

From the result image 9, we can clearly see that the edge segmentation method based on DCNN is smoother than the other three methods, and the classification is relatively concentrated, and the structure level is also clearer, especially the comparison method 1 (no ASPP is adopted). The method of structure), the accuracy of segmentation has been significantly improved. Based on this, we evaluate the segmentation accuracy and classification accuracy of the four segmentation methods.

C. SEGMENTATION PERFORMANCE EVALUATION

MRI is used to acquire brain information and create a 3D representation, but since the image lacks resolution in the third dimension, it is clipped along the direction of the feature axis to obtain a 2D slice. The convolution of the image is performed at the level of the image block, and 33×33 two-dimensional data blocks are extracted as training data from each pixel in the 2D slice for training. Manual segmentation marks the brain area as 5 parts (as shown in Figure 10): no tumor, edema, necrosis, non-sustained growth of the tumor, continuous growth of the tumor, we divide these five categories into three mutually concomitant tumor areas to The performance of the segmentation algorithm was evaluated, the Complete region (all four tumor frameworks), the Core region (necrosis, non-sustained growth tumors, persistently growing tumors), and the Enchancing region (continuously growing tumors). The evaluation indicators of the algorithm are as follows:

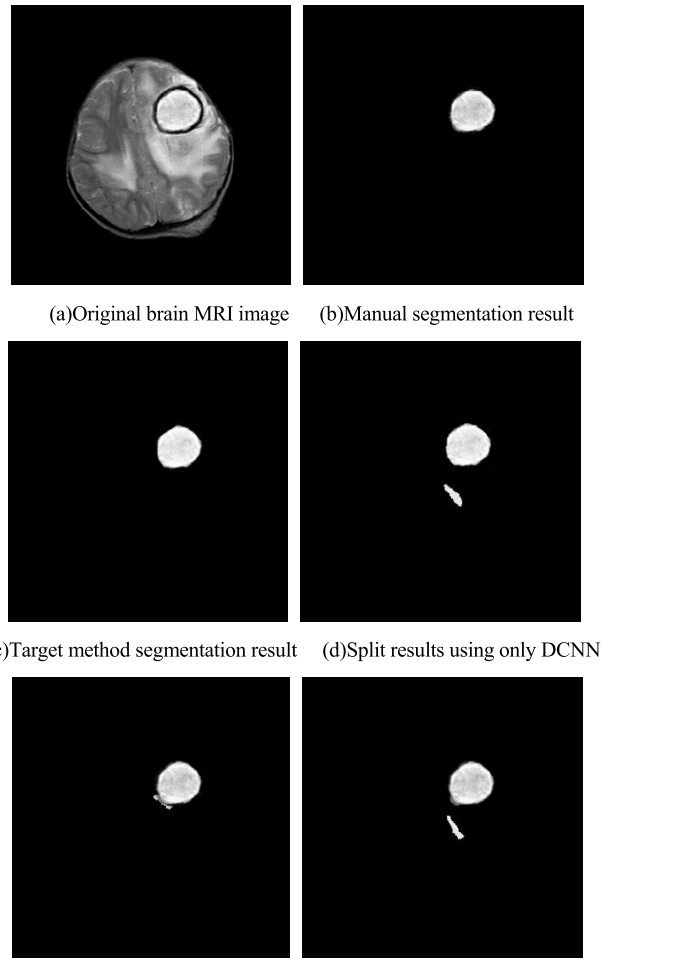
$$Dice = \frac{2TP}{2TP + FP + FN} \tag{12}$$

$$PPV = \frac{TN}{TN + FP} \tag{13}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{14}$$

Dice is used to measure the repeatability between the true value and the ideal value. PPV is the ratio of the segmentation of the correct tumor point to the segmentation result of all the segmentation into tumor points. Sensitivity is the ratio of the segmentation of the correct tumor point to the true tumor point. TP represents the number of true positives, FP represents the number of false positives, TN represents the number of true negatives, and FN represents the number of false negatives.

Four segmented MRI image samples were processed through research and evaluated. The evaluation coefficients for the final segmentation results obtained through the entire segmentation process are shown below. As can be seen from Table 2, the average value of the Dice coefficient is 0.86, and the variance is 0.0024; the average value of the positive



(a)Original brain MRI image (b)Manual segmentation result

(c)Target method segmentation result (d)Split results using only DCNN

(e)Comparison network 1 segmentation result (f)Comparison network 2 segmentation result

FIGURE 9. Experimental results of patient MRI images.

predictive value is 0.90, the variance is 0.0023; the average value of the sensitivity is 0.88, and the variance is 0.0011. This means that among the three evaluation methods, the values of the 12 randomly selected test data are relatively stable.

D. PERFORMANCE ANALYSIS OF THIS EXPERIM ENTAL COMPARISON METHOD

Combined with the experimental data in Table 3, the Dice coefficient of the experimental method is superior to the two control methods, indicating that the ASPP structure and CRF can indeed improve the performance of DCNN in MRI image segmentation of brain tumors. In particular, compared with the control method 1, the experimental method has obvious advantages, indicating that the ASPP structure has a significant influence on the improvement of the DCNN. In terms of positive detection values, the experimental method is not as effective as the control method in segmenting sub-tumor tissue or highlight regions. This may have something to do with the lack of training. Increasing the number of trainings, or modifying the parameters of the Adam method used to optimize parameters, may change this situation. In terms of sensitivity, the experimental method has a greater advantage.

TABLE 2. Total end-segmentation result evaluation coefficient.

Number	1	2
Dice coefficient	0.9288	0.7858
Positive predictive value	0.8697	0.9321
Sensitivity	0.9136	0.9077
3	4	5
0.8398	0.8685	0.8237
0.8062	0.9156	0.9269
0.8752	0.7963	0.8950
Number	6	7
Dice coefficient	0.8998	0.9353
Positive predictive value	0.8364	0.9007
Sensitivity	0.8995	0.8837
8	9	10
0.8934	0.8266	0.8085
0.9501	0.9642	0.8841
0.8558	0.8673	0.8981
11	12	13
0.8454	0.8527	0.9135
0.9369	0.8725	0.8943
0.8527	0.9053	0.8935

Compared with other control methods, this method was found to be at a high level of precision in MRI image segmentation of brain tumors. First of all, the deep convolutional

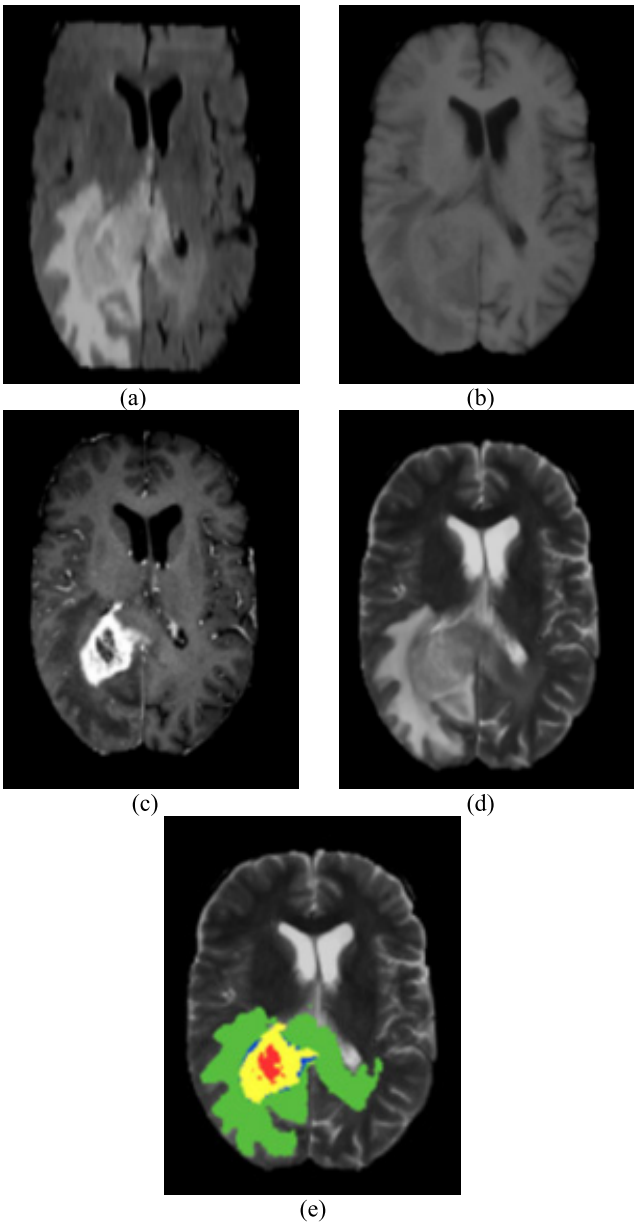


FIGURE 10. Four segmented MRI image samples (Colors are as follows: green - edema, blue - necrosis, yellow - non-sustainable growth of tumor, red - persistent growth of tumor).

TABLE 3. Comparison of performance of four segmentation methods.

Method	Dice
No ASPP structure	0.84
No CRF algorithm	0.85
Use only DCNN network	0.86
Target algorithm	0.86
Positive predictive value	Sensitivity
0.77	0.85
0.82	0.86
0.86	0.88
0.90	0.88

neural network is welcomed by researchers. The absolute number and quantity are increasing, the segmentation result is obviously better than other methods, and the network is continuously optimized as the research deepens. On the other hand, DCNN-based methods have several advantages over other convolutional neural networks. Compared to the two control networks, this advantage should be attributed to the ASPP structure and the use of CRF. At the same time, the method is more prominent in the segmentation of tumor sites, especially sub-tumors, which is closely related to the application of cascade structures.

VI. CONCLUSION

The combination of feature fusion and DCNN brain tumor MRI segmentation method proposed in this paper adopts cascade structure, which has significant improvement effect on segmentation of multi-level images. On the one hand, the complexity of a single network can be effectively reduced, and the contradiction between classification accuracy and segmentation accuracy can be solved. On the other hand, the original hierarchical structure of the image is used as a space constraint, which conforms to the laws of anatomy and physiology. Under the premise that the number of parameters is constant and the amount of calculation is constant, the ASPP structure subtly applies the cavity convolution to obtain the receptive fields of different sizes, and then fuses the information of different scales. The residual connection effectively avoids the gradient dispersion caused by the network being too deep. The CRF supplements the relationship between the pixels, and encourages pixels with similar intensity and position to be divided into the same area, and pixels with larger differences are divided into different areas. While using simpler operations, hierarchical, multi-scale, and high-quality segmentation, this method also has some shortcomings, that is, the method requires relatively low training time, so the number of iterations can only be minimized. The extent to which the accuracy of the segmentation results is affected. Therefore, the advantages of the experimental method may be more obvious after increasing the number of iterations. However, this is only an obstacle in the process of method research. In practical applications, it has sufficient good conditions and sufficient training time. Therefore, this limitation will not affect the feasibility of this method and meet the needs of practical applications. At the same time, future work will focus on better filters for processing brain tumor images, or combining features with better segmentation results, and how to further determine the types of intracranial tumors detected and segmented according to the tumor image, and help doctors More accurate types of regional tumors, diagnosis of the patient's condition, more intuitive clinical observation, and instead improve the cure rate of patients with brain tumors.

REFERENCES

[1] Z. Che, Y. Cheng, Z. Sun, and Y. Liu, "Exploiting convolutional neural network for risk prediction with medical feature embedding," 2017, *arXiv:1701.07474*. [Online]. Available: http://arxiv.org/abs/1701.07474

[2] M. Vardhana, N. Arunkumar, S. Lasrado, E. Abdulhay, and G. Ramirez-Gonzalez, "Convolutional neural network for bio-medical image segmentation with hardware acceleration," *Cognit. Syst. Res.*, vol. 50, pp. 10–14, Aug. 2018.

[3] S. Wei, W. Wu, G. Jeon, A. Ahmad, and X. Yang, "Improving resolution of medical images with deep dense convolutional neural network," *Concurrency Comput., Pract. Exper.*, vol. 32, no. 1, Jan. 2020, Art. no. e5084.
[4] Q. Li, W. Cai, and X. Wang, "Medical image classification with convolutional neural network," in *Proc. Int. Conf. Control Automat. Robot. Vis.*, Dec. 2014, pp. 844–848.
[5] X. X. LI and Z. Y. Jia, "Auxiliary detection method for small cell lung cancer based on convolutional neural network," *Chin. J. Digit. Med.*, vol. 8, no. 10, pp. 65–67, 2013.
[6] Y. Y. Wang, T. Zhou, and H. L. Lu, "Computer aided diagnosis model of lung tumor based on integrated convolutional neural network," *Biomed. Eng. J.*, vol. 34, no. 4, pp. 65–73, 2017.
[7] K.-J. Xia, H.-S. Yin, and J.-Q. Wang, "A novel improved deep convolutional neural network model for medical image fusion," *Cluster Comput.*, vol. 22, no. S1, pp. 1515–1527, Jan. 2019.
[8] S. D. Yu, "Research on convolutional neural network and migration learning in cancer image analysis," Ph.D. dissertation, Univ. Chin. Acad. Sci., Beijing, China, 2018.
[9] Y. Chen, "Research on automatic segmentation algorithm of cerebrovascular based on convolutional neural network," M.S. thesis, Univ. Electron. Sci. Technol. China, Sichuan, China, 2017.
[10] P. Fonseca, J. Mendoza, J. Wainer, J. Ferrer, J. Pinto, J. Guerrero, and B. Castaneda, "Automatic breast density classification using a convolutional neural network architecture search procedure," *Proc. SPIE Med. Imag., Comput.-Aided Diagnosis*, vol. 9414, Mar. 2015, Art. no. 941428.
[11] L. F. Zhai, J. Y. He, and M. Jian, "FCN-CNN cloud image segmentation method based on local cluster analysis," *J. Softw.*, vol. 29, no. 4, pp. 1049–1059, 2018.
[12] H. Lu, H. Wang, Q. Zhang, D. Won, and S. W. Yoon, "A dual-tree complex wavelet transform based convolutional neural network for human thyroid medical image segmentation," in *Proc. IEEE Int. Conf. Healthcare Informat. (ICHI)*, Jun. 2018, pp. 191–198.
[13] Q. S. Song, C. Zhang, and W. Chen, "Combined full convolutional neural network and conditional random field road segmentation," *J. Tsinghua Univ.*, vol. 58, no. 8, pp. 725–731, 2018.
[14] Z. Yang, Y. Huang, Y. Jiang, Y. Sun, Y.-J. Zhang, and P. Luo, "Clinical assistant diagnosis for electronic medical record based on convolutional neural network," *Sci. Rep.*, vol. 8, no. 1, Dec. 2018, Art. no. 6329.
[15] A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical image retrieval using deep convolutional neural network," *Neurocomputing*, vol. 266, pp. 8–20, Nov. 2017.
[16] S. Miao, Z. J. Wang, and R. Liao, "A convolutional neural network approach for 2D/3D medical image registration," *Optoelectron. Adv. Mater.-Rapid Commun.*, vol. 4, no. 3, pp. 636–648, 2015.
[17] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.
[18] W. N. Wang, L. Wang, and M. Q. Zhao, "Image Aesthetic Classification Based on Parallel Deep Convolutional Neural Network," *Acta Automatica Sinica.*, vol. 42, no. 6, pp. 904–914, 2016.
[19] A. Yang, Y. Li, C. Liu, J. Li, Y. Zhang, and J. Wang, "Research on logistics supply chain of iron and steel enterprises based on block chain technology," *Future Gener. Comput. Syst.*, vol. 101, pp. 635–645, Dec. 2019.
[20] J. Li, L. Zhang, X. Feng, K. Jia, and F. Kong, "Feature extraction and area identification of wireless channel in mobile communication," *J. Internet Technol.*, vol. 20, no. 2, pp. 545–553, 2019.
[21] G. C. Zhu and C. F. Li, "Detection of chest and lung nodules based on convolutional neural network," *J. Sensors Microsyst.*, vol. 36, no. 12, pp. 153–156, 2017.
[22] T. Zheng, Y. Gao, F. Wang, C. Fan, X. Fu, M. Li, Y. Zhang, S. Zhang, and H. Ma, "Detection of medical text semantic similarity based on convolutional neural network," *BMC Med. Informat. Decis. Making*, vol. 19, no. 1, p. 156, Dec. 2019.
[23] K. H. Cha, L. Hadjiiski, R. K. Samala, H.-P. Chan, E. M. Caoili, and R. H. Cohan, "Urinary bladder segmentation in CT urography using deep-learning convolutional neural network and level sets," *Med. Phys.*, vol. 43, no. 4, pp. 1882–1896, Mar. 2016.
[24] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1240–1251, May 2016.
[25] X. L. Yang, Y. Qiang, and W. Zhao, "CT image hash retrieval of pulmonary nodules based on medical signs and convolutional neural networks," *J. Intell. Syst.*, vol. 12, no. 6, pp. 857–864.2017.
[26] J. M. Wolterink, T. Leiner, B. D. de Vos, R. W. van Hamersvelt, M. A. Viergever, and I. Išgum, "Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks," *Med. Image Anal.*, vol. 34, pp. 123–136, Dec. 2016.
[27] A. Yang, C. Zhang, Y. Chen, Y. Zhuansun, and H. Liu, "Security and privacy of smart home systems based on the Internet of Things and stereo matching algorithms," *IEEE Internet Things J.*, early access, Oct. 8, 2019, doi: 10.1109/JIOT.2019.2946214.
[28] W. Jifara, F. Jiang, S. Rho, M. Cheng, and S. Liu, "Medical image denoising using convolutional neural network: A residual learning approach," *J. Supercomput.*, vol. 75, no. 2, pp. 704–718, Feb. 2019.
[29] H. Fu, G. Manogaran, K. Wu, M. Cao, S. Jiang, and A. Yang, "Intelligent decision-making of online shopping behavior based on Internet of Things," *Int. J. Inf. Manage.*, vol. 50, pp. 515–525, Feb. 2020.
[30] Z. Li, N. Dey, A. Ashour, L. Cao, Y. Wang, D. Wang, P. McCauley, V. Balas, K. Shi, and F. Shi, "Convolutional neural network based clustering and manifold learning method for diabetic plantar pressure imaging dataset," *J. Med. Imag. Health Informat.*, vol. 7, no. 3, pp. 639–652, Jun. 2017.

NAIQIN FENG was born in Xinxiang, Henan, China, in 1953. He received the B.S. degree in communication engineering from the Nanjing Institute of Communication Engineering, Nanjing, in 1977, and the Ph.D. degree in basic psychology (research direction: artificial intelligence) from Southwest University, Chongqing, in 2007.

From 1986 to 1994, he was an Engineer with the Computing Center, Henan Normal University, where he was promoted to a Senior Engineer, in 1995. Since 2008, he has been a Professor with the College of Computer and Information Engineering, Henan Normal University. He is the author of three books and more than 50 articles. He has participated in two national projects and presided over six provincial scientific research projects. His research interests include artificial intelligence, artificial neural networks, associative memory, and fuzzy theory.

Dr. Feng is an Invited Editor of the *Software* journal.

XIUQIN GENG was born in Xinxiang, Henan, China, in 1961. She received the B.S. degree in medicine from the Henan Medical College, Zhengzhou, in 1983.

From 1984 to 1990, she was a Resident Physician with the Internal Medicine Department, Xinxiang Central Hospital. In December 1991, she was an Attending Physician; in December 1996, she was the Deputy Chief Physician; in December 2003, she was the Chief Physician with the Xinxiang Central Hospital; and a part-time Professor and a Master Tutor with the Medical School, Zhengzhou University, and Xinxiang Medical School. She is the author of three books and more than 30 articles. She has participated in two national projects and presided over three provincial scientific research projects. Her research interests include endocrine and metabolic diseases, complications of diabetes, thyroid disease, growth retardation, computer assisted bone age evaluation, and computer application in medicine.

LIJUAN QIN was born in Zhoukou, Henan, China, in 1986. She received the B.S. degree in computer science and technology and the M.S. degree in computer software and theory from Henan Normal University, Xinxiang, China, in 2009 and 2013, respectively.

From 2015 to 2018, she has worked as an Assistant Professor with the School of Information Engineering, Zhengzhou University of Industrial Technology. In 2019, she has been a Lecturer of computer science and technology. She has written a book and six articles and participated in one national project and three provincial and municipal scientific research projects. Her research interests include morphological neural networks, morphological associative memory, and pattern recognition.