Neurocomputing 350 (2019) 117–127

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Efficient multi-kernel DCNN with pixel dropout for stroke MRI segmentation

Liangliang Liu[a], Fang-Xiang Wu[b], Jianxin Wang[a,*]

[a] School of Computer Science and Engineering, Central South University, Changsha 410083, PR China
[b] Department of Mechanical Engineering and Division of Biomedical Engineering, University of Saskatchewan, Saskatoon, SK S7N5A9, Canada

ARTICLE INFO

Article history:
Received 5 January 2019
Revised 4 March 2019
Accepted 22 March 2019
Available online 18 April 2019

Communicated by Dr Xiaofeng Zhu

Keywords:
Magnetic resonance imaging
Deep convolution neural network
Dense block
Multi-kernel
Dropout

ABSTRACT

As manual delineation of lesions in medical image is a very tedious and time consuming process, accurate and automatic segmentation of medical images can assist diagnosis and treatment. In this study, we propose a deep convolution neural network for stroke magnetic resonance imaging(MRI) segmentation. The main structure of our network consists of two symmetrical deep sub-networks, in which dense blocks are embedded for extracting effective features from sparse pixels to alleviate the over-fitting problem of deep networks. We use the multi-kernel to divide the network into two sub-networks for acquiring more receptive fields, and the dropout regularization method to achieve an effective feature mapping. For the post-processing of the soft segmentation, we use image median filtering to alleviate noises and preserve the edge details of images. Our network is evaluated on two public benchmark segmentation challenges (SISS: sub-acute ischemic stroke lesion segmentation and SPES: acute stroke outcome/penumbra estimation) with multi-modality MRI sequences. According to the results of the public benchmark reports, among 9 teams participating in both SISS and SPES challenges at the same time, our network achieves the top performance on SISS challenge, and the top 3 performance on the SPES challenge. In addition, our network also exhibits state-of-the-art performance compared with other segmentation methods. Finally, we extensively evaluate our network with an ablation experiment. The experimental results show that both multi-kernel and dropout strategies can improve the segmentation accuracy of our proposed network.

1. Introduction

Accurate segmentation of the lesions can provide vital information for subsequent quantitative prediction of disease and treatment strategies. In clinical diagnosis, segmentation of lesion tissue is accomplished manually in medical images. The diversity of medical images, the variability of lesions and the large number of image slices demand a huge amount of time for neurologists to make decision, and the quality of manual segmentation is directly related to their states [1]. More importantly, neurologists usually incorporate their own experiences into the image segmentation. However, how to translate these experiences to automatic image segmentation by computer algorithm is non trivial. In addition, in medical images, there are generally no obvious boundaries between lesions and the surrounding tissues, which impedes the performance of segmentation methods. On the other hand, the similar signals may represent different diseases or indistinguishable noises [2,3], which are challenging for medical image segmentation. In order to alleviate the above issues, some unsupervised methods are used in medical image studies. Cardoso et al. proposed a mixture-model for imaging synthesis which was used to locate pathological regions [4]. Forbes et al. presented a Bayesian multi-sequence Markov model for detecting multiple sclerosis and stroke lesions [5]. Erihov et al. extended an image saliency algorithm to handle tumor detection based on medical images [6]. These methods do not require the manual ground truth. However, these methods focus on the detection of lesions rather than the segmentation of lesions.

Recent years, automatic segmentation of lesions is a major focus of medical image studies. Deep learning has made great achievements in the field of artificial intelligence in the past several years [7]. Especially in the imaging applications, deep learning provides a powerful tool for medical image segmentation [8–10] and medical image classification [11–13]. Deep convolution neural network (DCNN) is one of main tools to deal with image problems [14]. It has been successful in Computed Tomography (CT) scan \Magnetic Resonance Imaging (MRI) image segmentation and

* Corresponding author.
E-mail address: jxwang@mail.csu.edu.cn (J. Wang).

https://doi.org/10.1016/j.neucom.2019.03.049
0925-2312/

disease diagnosis [13,15–18]. In this study, we are interested in the automatic segmentation of lesions from stroke MRI sequences using an efficient DCNN.

The rest of this paper is structured as follows. We first introduce some stroke lesions segmentation related work in Section 2. Then we introduce the architecture of the proposed network in Section 3. We provide a quantitative evaluation of our segmentation network on SISS and SPES challenges in Section 4. We provide ablation experiments to quantify the main contributions of individual components of our proposed network in Section 5. Finally, we discuss and summarize this study in Sections 6 and 7, respectively.

2. Related work

Stroke is the second cause of disability and death, and thus correct and rapid assessment of the presence and location of stroke lesions is very important for the treatment of the disease. CT and MRI have always been the main information carriers in non-invasive stroke research. CT has been used to triage stroke lesions, mainly because of its accessibility, speed, availability and lack of contraindications. In the past decades, many methods for stroke segmentation of CT have been proposed. Maldjian et al. presented a method for detecting the lesions of acute ischemia stroke, which is based on CT images [19]. They compared the voxel density in the lenticular nucleus and the insular lobes with the contralateral side, and judged whether there was stroke infarct in the detection regions by the density. Usinskas et al. proposed a simple unsupervised segmentation method to segment ischemic stroke lesions on CT images [20]. This method is based on the 18 joint texture features determined by neurologists. They used mean, standard deviation, histogram and gray level co-occurrence matrix methods to identify the ischemic stroke lesions. Poh et al. proposed an automated template-based method to segment ventricular cerebrospinal fluid lesions in stroke CT images [21]. They used the thresholding method to divide the region of interest and accounted for the presence of stroke lesions. Gillebert et al. proposed a method for automatical segmentation of infarct and hemorrhage in stroke CT images, which combined the template-based method with the regional contrast method [22]. Although these methods reduce the manual rendering time, their sensitivity is too low. In addition, CT images are not the preferred choice for stroke diagnose. CT images are obtained by plain scan, which can not identify the tissue with the poor density and the low resolution.

MRI can scan transversely and longitudinally, and has better resolution in soft tissue. Many studies have been dedicated to the automatic segmentation of stroke based on MRIs, including region growing algorithm, Markov random field, thresholding method, random forest (RF), and so on. The region growing algorithm is a region-based image segmentation method, which requires initial seed points and uses an iterative method to determine whether the pixel neighbors belong to the same region. Saad et al. used a region growing algorithm to detect the lesion regions [23]. They used the difference of pixel strength as a measure of segmentation. Kabir et al. proposed a multimodal Markov random field model for stroke lesions segmentation on MRI sequences. The model is constructed on the Atlas of blood supply territories for assisting diagnosis [24]. Mitraa et al. presented an approach to identify ischemic lesions [25]. First, they used Bayesian Markov random field to classify the lesion tissues, then they used RF to extract areas with high likelihood. In the end the thresholding method was used to obtain the segmentation result. An important drawback of these methods is that they mainly use the hand-crafted feature.

More recently, convolutional neural networks (CNNs) have achieved great success in the field of artificial intelligence, including image classification, detection and recognition [26–28]. CNNs can learn features from raw images and extract context information. The feature sets, which are filtered by CNNs, often outperform pre-defined and hand-crafted feature sets. Many extended CNN methods were used in medical image segmentation tasks. For example, a multi-channel fully convolutional network (FCN) was used to segment liver tumors [29]. Fully convolutional residual networks (FC-ResNets) was used in electron microscopy image segmentation [30]. One 11-layer CNN with the dual pathway was used for brain lesion segmentation [31]. A U-shape model was used for biomedical image segmentation [15]. Most of the top ranking methods in ISLES 2015 stroke segmentation challenge are based on CNNs [18]. In addition, DCNNs have gained much popularity and made great achievements in the medical image segmentation. Compared with traditional CNNs, these DCNN models have more layers or structural changes. Liu et al. proposed a new network (Res-FCN), which combined the residual block with FCN. This network was used to automatically segment ischemic stroke lesions from MRIs [32]. Similarly, Liu et al. proposed a network (Res-CNN), which combined the residual block with U-shape network [33]. In addition, Guerrero et al. proposed a DCNN for detecting WMH and stroke, which embedded residual units into U-shape network for obtaining more context features [34]. This network not only could segment the WMH lesions, but also could differentiate WMH and stroke lesions. Zhang et al. proposed a DCNN for acute ischemic stroke segmentation, which combined the FCN and dense blocks [35]. The dense blocks not only could relieve the difficulty of training DCNN, but also could get more context information without adding training parameters. Though DCNNs are powerful, there are some deficiencies that discouraged the employment of DCNNs on medical imaging data. The deeper a network, the more hyper-parameters are needed for training it. It is notorious that deep networks cause serious vanishing-gradient [36] and over-fitting problems [37].

In this study, in order to alleviate the above deficiencies of DCNNs. We propose a novel end-to-end multi-kernel DCNN (MK-DCNN) for stroke MRI segmentation. We investigate the performance of our network across two different public challenges with different scanners and acquisition protocols. The main contributions of our study are as follows:

1. We propose an end-to-end DCNN which is composed of U-shape architecture and dense blocks. The U-shape uses the underlying features to improve the lack of information on the sampling. Dense blocks are committed to alleviating the vanishing-gradient problem and improving the network performance from depth and breadth.
2. We divide the network into two sub-networks by two different convolution kernels in the first layer. This strategy can extract more image features than the single kernel. We combine the two sub-networks before output. This design strategy can help improve the performance of the segmentation.
3. We embed the dropout regularization method in dense blocks and transition blocks. It's an effective method to prevent neural networks from the over-fitting problem [37,38]. In our experiment, the dropout regularization method can improve the accuracy of segmentation.
4. Our network is superior to the most advanced technology on two sub challenges of the ISLES in 2015. In addition, we conduct ablation experiments to analyze the performance of the network.

3. Method

The MK-DCNN framework we propose in this study, is mainly based on the U-shape architecture. The U-shape is an excellent architecture for image segmentation tasks with small samples, which outperforms the state-of-the-art on a variety of medical image segmentation tasks [15,17,39–41]. The U-shape architecture consists of

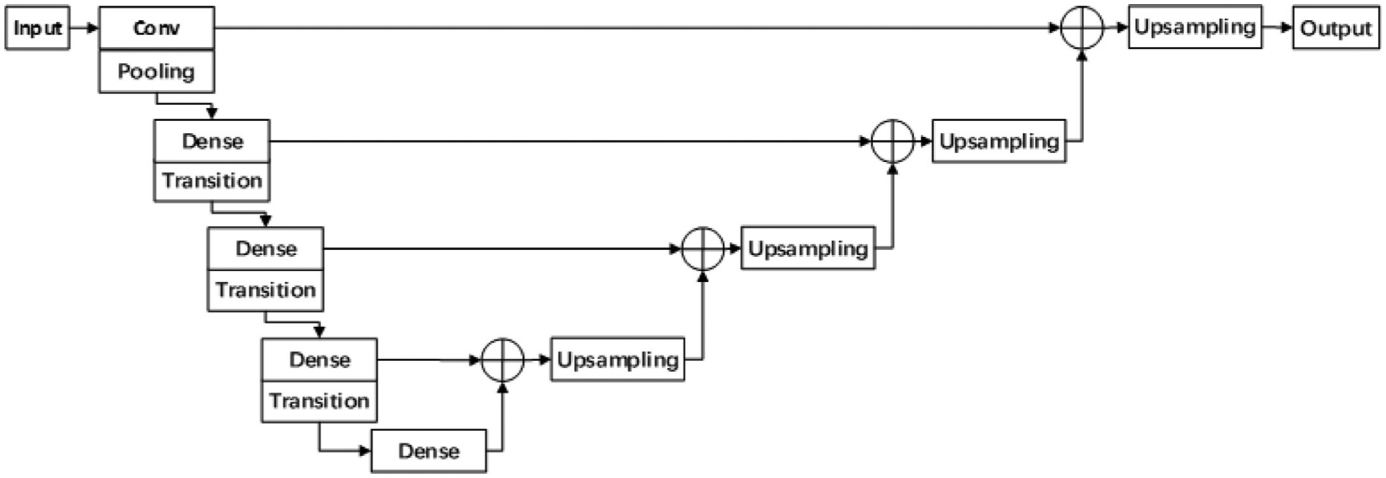


Fig. 1. Architecture of the sub-network in MK-DCNN.

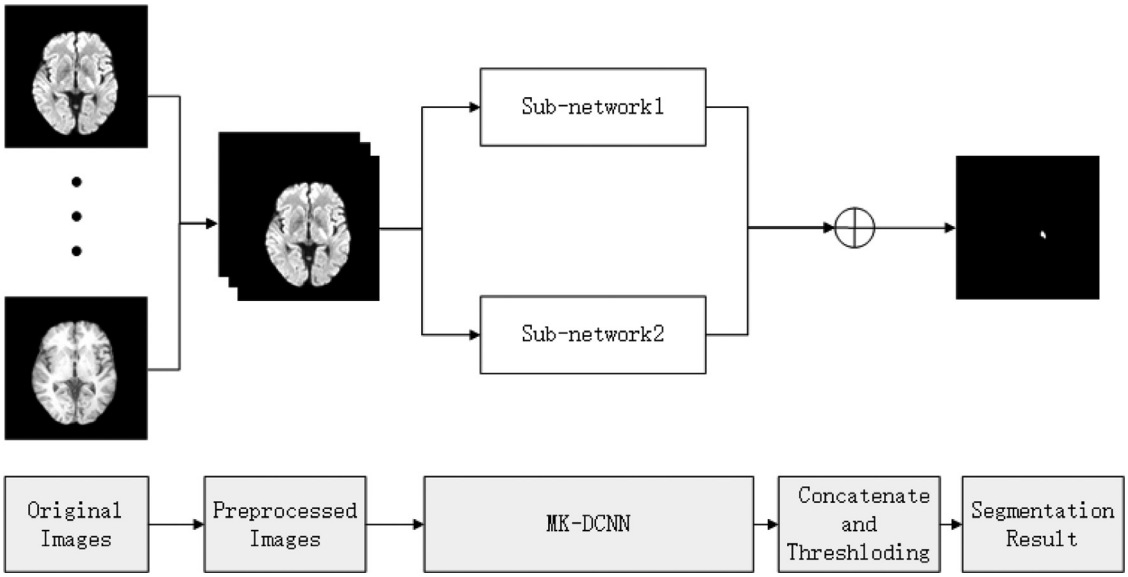


Fig. 2. The pipline of the MK-DCNN.

two paths: a contracting path and a expanding path. The contracting path is used for capturing context and the expanding path is used for precise positioning. We embed the dense structure as a block into the contracting path of the U-shape. The dense block not only reduces training parameters or alleviates the vanishing-gradient problem, but also realizes the reuse of features [42,43].

3.1. Analysis of network architecture

MK-DCNN consists of two sub-networks. Fig. 1 illustrates the architecture of our proposed sub-network, which is an end-to-end deep neural network. This architecture inherits both of advantages of U-net and DenseNet [15,43]. The external framework of the sub-network is mainly based on U-shapes, and the internal feature extraction pipeline architecture follows DenseNet-121 [43]. In the contracting path of the U-shape, we uses the dense blocks and transition blocks to replace the convolution layers and pooling layers in the traditional U-net. The dense block is used to alleviate the vanishing-gradient problem, encourage feature reuse, and substantially reduce the number of parameters. The transition block is used to complete the operation of pooling and reduce the size of the next input layer. The original DenseNet-121 is used to object classification [43]. However, we want to implement the task of image segmentation. Therefore, we keep the successive up-sampling layers of the original U-net, which are used in the expanding path to the precise position of pixels. In the middle of each sub-network, we connect the peer-to-peer layer to crop feature map from the contracting path to the expanding path, which can help get more feature information.

As shown in Fig. 2, MK-DCNN consists of two symmetrical and parallel sub-networks. The architecture of each sub-network is similar to Fig. 1. The preprocessed images are send into MK-DCNN, in the first layer, two convolutions with different kernel sizes lead the images to two sub-networks. In the end, the output feature maps of two sub-networks, which are combined into one feature map by concatenation operation. Then the final segmentation result is generated.

The details of the network structure are shown in Table 1. MK-DCNN has two sub-networks. Each sub-network consists of 3 convolution layers, 1 pooling layer, 4 dense blocks, 3 transition blocks and 4 up-sampling blocks. After the image is input into a sub-network, we first use convolution and pooling layers to complete the preliminary features extraction. Then important features extraction and down-sampling operation is completed by dense

Table 1
Architecture of MK-DCNN. The two sub columns in the second column represent two sub-networks. The "concat" means Concatenate. Note that each "conv" consists of the sequence BN-ReLU-Conv.

Block	Architecture	
Input	Images	
Convolution	Conv A:3×3	Conv B:7×7
Pooling	2×2 Maxpool, stride 2	2×2 Maxpool, stride 2
Dense Block	Dense A1: $\begin{bmatrix} 1 \times 1, 128 & conv \\ 3 \times 3, 32 & conv \end{bmatrix} \times 6$	Dense B1: $\begin{bmatrix} 1 \times 1, 192 & conv \\ 3 \times 3, 48 & conv \end{bmatrix} \times 6$
Transition Block	Transition: $\begin{bmatrix} 1 \times 1 & conv \\ 2 \times 2, avgpool & stride\ 2 \end{bmatrix}$	Transition: $\begin{bmatrix} 1 \times 1 & conv \\ 2 \times 2, avgpool & stride\ 2 \end{bmatrix}$
Dense Block	Dense A2: $\begin{bmatrix} 1 \times 1, 128 & conv \\ 3 \times 3, 32 & conv \end{bmatrix} \times 12$	Dense B2: $\begin{bmatrix} 1 \times 1, 192 & conv \\ 3 \times 3, 48 & conv \end{bmatrix} \times 12$
Transition Block	Transition: $\begin{bmatrix} 1 \times 1 & conv \\ 2 \times 2, avgpool & stride\ 2 \end{bmatrix}$	Transition: $\begin{bmatrix} 1 \times 1 & conv \\ 2 \times 2, avgpool & stride\ 2 \end{bmatrix}$
Dense Block	Dense A3: $\begin{bmatrix} 1 \times 1, 128 & conv \\ 3 \times 3, 32 & conv \end{bmatrix} \times 24$	Dense B3: $\begin{bmatrix} 1 \times 1, 192 & conv \\ 3 \times 3, 48 & conv \end{bmatrix} \times 24$
Transition Block	Transition: $\begin{bmatrix} 1 \times 1 & conv \\ 2 \times 2, avgpool & stride\ 2 \end{bmatrix}$	Transition: $\begin{bmatrix} 1 \times 1 & conv \\ 2 \times 2, avgpool & stride\ 2 \end{bmatrix}$
Dense Block	Dense A4: $\begin{bmatrix} 1 \times 1, 128 & conv \\ 3 \times 3, 32 & conv \end{bmatrix} \times 16$	Dense B4: $\begin{bmatrix} 1 \times 1, 192 & conv \\ 3 \times 3, 48 & conv \end{bmatrix} \times 16$
Upsampling Block	Upsampling A1: 2×2 up-sampling, concat[Upsampling A1,Dense A3],504,conv	Upsampling B1: 2×2 up-sampling, concat[Upsampling B1,Dense B3],768,conv
Upsampling Block	Upsampling A2: 2×2 up-sampling, concat[Upsampling A2,Dense A2],224,conv	Upsampling B2: 2×2 up-sampling, concat[Upsampling B2,Dense B2],384,conv
Upsampling Block	Upsampling A3: 2×2 up-sampling, concat[Upsampling A3,Dense A1],192,conv	Upsampling B3: 2×2 up-sampling, concat[Upsampling B3,Dense B1],96,conv
Upsampling Block	Upsampling A4: 2×2 up-sampling, concat[Upsampling A4,Conv A],96,conv	Upsampling B4: 2×2 up-sampling, concat[Upsampling B4,Conv B],96,conv
Concatenate	concat[Upsampling A4,Upsampling B4], BN/ReLU	
Convolution	Conv C: 2×2 up-sampling, 1×1 conv	
Convolution	Conv D:1×1,conv sigmoid	
Output	Lesion segmentation	

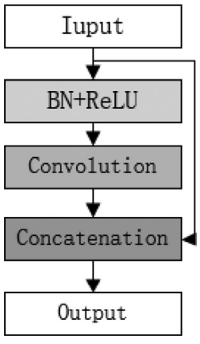


Fig. 3. Architecture of a dense micro-unit.

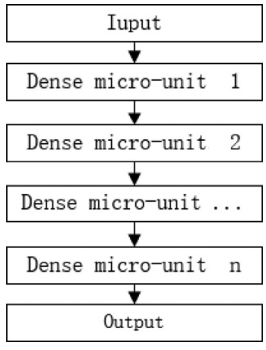


Fig. 4. Architecture of the n-layer dense block.

blocks and transition blocks. As shown in Fig. 4, there is a dense block which is used to alleviate the vanishing-gradient problem and improve the network performance from depth and breadth. A dense block is consist of n-layer dense micro-units ($n \geq 1$). As shown in Fig. 3, a dense micro-unit consists of a batch normalization (BN) [36] layer, a rectified linear unit (ReLU) [44] layer and a convolution (Conv) layer. The structure of transition block is shown Fig. 5. A transition block consists of a BN layer, a ReLU layer and an average pooling layer [14]. Transition block is used to improve the speed of network training and realize down-sampling. Before generating the segmentation results, we use upsampling blocks enlarge the dimension of image. Finally, after two convolution layers, we get a feature map from sub-network. Then we combined two feature maps into one by concatenation operation and use the sigmoid function to complete the segmentation task.

3.2. Feature mapping

In medical image segmentation tasks, an image is a pixel matrix, and a pixel is a coordinate in the matrix. Each pixel label is independent of its neighborhood in an image. MK-DCNN predicts every pixel-wise segmentation by coordinates and contextual information of pixels. It realizes feature mapping through convolution with multiple filters between adjacent layers. Finally, according to

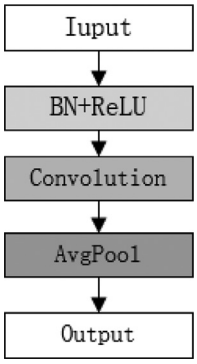


Fig. 5. Architecture of the transition block.

the activation function, we predict the type of each pixel and generate the segmentation result. Let X be the number of input layers, $x \in [1, X]$ denotes an input layer, FM_x be the feature mapping of the xth layer, n be the feature mapping number of the xth layer. The feature mapping obtained by convolution from previous layer $x-1$ can be described as follows:

$$y_x = f\left(\sum_{n=1}^{FM_{x-1}} k_x^n * y_{x-1}^n + \Theta_x\right), \tag{1}$$

where y_x is the result of convolving the $(x-1)$th layer with k_x^n. k_x^n is the kernel of the xth layer, which is also the hidden weight matrix (W_x^n) from layer $x-1$ to layer x. Θ_x is a learned bias at layer x. $f()$ is the non-linearity feature mapping function. If $x = 1$, y_0^n denotes the original input layer.

3.3. Parallel multi-kernel for feature exploration

Each pixel feature in the image determines the result of the segmentation more or less. In neural network, the convolution kernel is used to filter the contextual information of pixels from input images. If the parameters of the convolution kernel are fixed, the extracted feature information would be limited, which may easily lead to the deviation of the lesion segmentation. The receptive field of the convolution kernel is also limited, which may only focus on local features. Different kernels generate different receptive fields, and they can incorporate more features from the same image and provide multiple features for lesion analysis.

In MK-DCNN, we use dual different convolution kernels in the first layer to obtain different contextual information. The details of the kernels are shown in Table 1 (e.g. 3×3 and 7×7). An input image is sent into two sub-networks with different contextual information in the first convolution layer. The output image size is described as:

$$\begin{cases} height\_out = \left\lceil \dfrac{(height\_in - height\_kernel + 2 * padding)}{stride} \right\rceil + 1, \\ width\_out = \left\lceil \dfrac{(width\_in - width\_kernel + 2 * padding)}{stride} \right\rceil + 1, \end{cases} \tag{2}$$

where *height_out* and *width_out* denote the size of an output image, *height_in* and *width_in* denote the size of an input image. *height_kernel* and *width_kernel* denote the size of convolution kernel. *padding* is used to maintain boundary information and *stride* represents the step length. In the first convolution layer of our network, the *padding* is "same" and the *stride* is 1.

After filtered by two convolution kernels, different contextual information is sent to two contracting sub-networks for the down-sampling operation. The two sub-networks are symmetrical and independent. In the process of down-sampling, different context information is extracted by dense blocks and transition blocks. Up-sampling is used to restore image information. After down-sampling and up-sampling, the input images are transformed into two matrices with the same scale. The output matrices come from two sub-networks, which are combined into a matrix along an existing axis by concatenation operation, as shown in Fig. 6. Combining the context information from different networks can improve the performance of segmentation [15,45]. The multi-kernel provides more nest details with different receptive fields than one kernel. MK-DCNN can provide more alternative contextual features for segmentation.

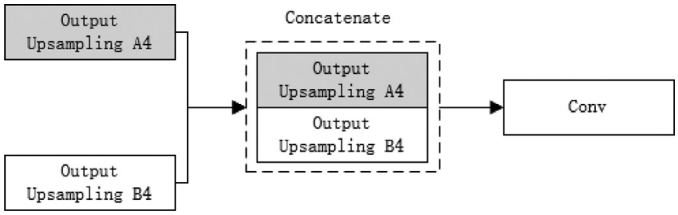


Fig. 6. Pipeline of the concatenate.

3.4. Dropout regularization method for effective learning

Multi-scale or multi-modality images can provide richer contextual feature information for segmentation model, which are benefit to boost the performance of models [31,46–49]. However, simply using multi-scale or multi-modality is a burden to DCNNs, which not only demands much training time, but also increases noise and degrades performance. How to find a balance between richer context features and important feature screening still needs to be explored.

The regularization methods are used to prevent the over-fitting problem, reduce the number of feature vectors and reduce the complexity of the model. These methods can automatically weaken the unimportant feature variables and extract important feature variables from many feature variables. Dropout is one of regularization methods, which is used to screen important features in hidden layers of the network and alleviate the over-fitting problem [37,38,50,51]. Compared with regularization methods L1 and L2, dropout does not aim at optimizing the cost function while it changes the structure of the neural network [37,52]. The essence of dropout is to restrict the parameters to be optimized [50]. In the DCNNs training process, dropout temporarily discards the hidden neurons according to a certain probability, which can preserve more robust and valuable hidden neurons in neural networks.

In MK-DCNN, we use the multi-kernel method to generates more contextual feature information. However, different contextual feature information may relate to the same location of the image. It means that multiple neurons in the network may detect the same feature, which wastes resources. In order to alleviate this problem, in MK-DCNN, we use the dropout regularization method to filter feature information. The application of dropout on a generic ith neuron is shown below:

$$Q_i = x_i a\left(\sum_{k=1}^{d_i} w_k x_k + b_k\right)(0 \le i \le h), \tag{3}$$

where Q_i is the retained probability of neuron x_i (a Bernoulli random variable), $a()$ is an activation function, $k \in [1, i]$ is the unit number, w_k and b_k are the kth unit weight and bias, respectively. d denotes dimensions, x_{di} indicates that x_i is a Bernoulli variables with d dimensions. $\sum_{k-1}^{d_i} w_k x_k$ is the sum of the product of all neuron weights w_k and x_k before the ith neuron.

In MK-DCNN, we need to dropout a set of neurons from a layer. Let the jth layer have n neurons. In a cycle, the neural network can be regarded as the integration n times of Bernoulli's experiments. Thus, the number of neurons retained in layer j can be computed as follows:

$$Y = \sum_{i=1}^{d_j} x_i, \tag{4}$$

where Y is the retained number of neurons, x_i is a retained neuron. In the n experiments, the probability of retaining k neurons was:

$$f(k; n, p) = \binom{n}{k} p^k q^{(n-k)}, \tag{5}$$

where $p = 1 - q$, p represents the probability of a neuron being kept on and q represents the probability of a neuron being turned off. $p^k q^{(n-k)}$ is the probability of obtaining a single sequence of k successes on n trials and $(n - k)$ failures, while $\binom{n}{k}$ is the binomial coefficient used to calculate the number of possible successful sequences.

In MK-DCNN, we use a fixed dropout ratio to handle the feature filtering in each training iteration. When a limited amount of neurons disappear at a fixed ratio in the hidden layers, a certain number of redundant features can be filtered out. In the end, the feature learning process can be enhanced in hidden layers of the network.

3.5. Loss function

Image is composed of pixels. Gray image segmentation task can be regarded as pixel's binary classification. Dice coefficient (DC) is one of the classic metrics for evaluating the segmentation performance. It can be used as loss function to measure the gap between the result of the segmentation and the ground truth [53]. Considering two sets of X and Y, the DC is defined as follows:

$$DC(X, Y) = \frac{2|X \bigcap Y|}{|X| + |Y|}, \tag{6}$$

where X and Y denote the set of ground truth and segmentation, respectively. $DC(X, Y) \in [0, 1]$, 0 indicates two sets have no overlap, and 1 indicates two sets are the exact same. In binary image segmentation, we use the softmax function output to replace the predicted binary labels. In this study, we combine DC with cross entropy function, and a pseudo DC loss function is defined as:

$$L = 1 - \frac{1}{C} \sum_{c=1}^{C} \left(\frac{2 \sum_{n=1}^{N} (p(x_n)^c q(x_n)^c)}{\sum_{n=1}^{N} q(x_n)^c + \sum_{n=1}^{N} p(x_n)^c} \right), \tag{7}$$

where C is the number of classes, $c \in C$ is the pixel class, N is the pixel number, x_n is the n-th pixel. $p(x_n)^c$ represents the true probability of pixel x_n belonging to class c, and $q(x_n)^c$ represents the prediction probability of pixel x_n belonging to class c. In order to measure the loss contribution of each class, the aggregating DC is defined as the average of DCs from different classes. In the traditional single type lesion segmentation task, C is usually set to 1.

Adam optimization algorithm [54] is outstanding in computer vision and natural language processing. It needs less resources and makes the model converge quickly. It also can accelerate the training speed fundamentally. In our experiment, we use Adam optimization algorithm to train MK-DCNN. In the prediction stage, the forecasting process of lesion is automatic in MK-DCNN without any manual intervention.

3.6. Evaluation measures

In our study, we adopt three metrics to evaluate the quality between the segmentation result and the reference ground-truth: the Dice's coefficient (DC), the Average Symmetric Surface Distance (ASSD) and the Hausdorff distance (HD). DC has been defined in Eq. 6, which measures the similarity between two images. ASSD denotes the average distance between the volumes surface points averaged over both directions. Considering two sets of surface points X and Y, the ASSD is defined as:

$$ASSD(X, Y) = \frac{\frac{\sum_{x \in X} min_{y \in Y} d(x,y)}{|X|} + \frac{\sum_{y \in Y} min_{x \in X} d(y,x)}{|Y|}}{2}, \tag{8}$$

where $\frac{\sum_{x \in X} min_{y \in Y} d(x,y)}{|X|}$ and $\frac{\sum_{y \in Y} min_{x \in X} d(y,x)}{|Y|}$ are the average surface distance base on $d(x, y)$ and $d(y, x)$, respectively. $d(x, y)$ being the Euclidean distance between the points x and y, and $d(y, x)$ being the Euclidean distance between the points y and x. ASSD is given in mm, the lower the better.

HD denotes the maximum distance between two volumes surface points. HD can denote outliers. It is defined as:

$$HD(X, Y) = \max \left\{ \max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(y, x) \right\}. \tag{9}$$

Similar to the ASSD, the HD is given in mm and a lower value denotes a better segmentation.

4. Experiments & results

MK-DCNN is applied in ISLES 2015 challenge. ISLES challenge has two different sub-tasks with MRIs: SISS and SPES. Two tasks are about ischemic stroke disease with different lesions and multi-modality MRIs. All participants verify and adjust the algorithms based on the benchmark training dataset. Finally, the testing dataset without the ground truth was distributed on the challenge web pages. Participants should submit their final segmentation results to the organizers, who scored the segmentation results.

4.1. SISS

Sub-acute ischemic stroke is a common cerebrovascular disease, which is a sub type of stroke disease. From the onset, this disease usually appears in the stage of 24H-2W [55]. There are 64 sub-acute ischemic stroke samples in the SISS dataset, 28 samples for training and 36 samples for testing, 56 out of which come from a medical center and 8 additional samples come from another neuroradiology center. All samples were anonymous patients who were diagnosed as ischemic stroke, and each sample has T1-w, T2-w, DWI and Flair MRI sequences. The purpose of the SISS challenge is to segment the lesions of sub-acute infarct. All sequences are already skull-stripped and co-registered by the organizers. The samples in both training and testing datasets are preserved the diversity of the stroke cases: both contain single- and multi-focal, small and large lesions. In our experiment, we use multi-modality MRIs (DWI and Flair) as inputs.

4.2. SPES

Penumbra is surrounding tissue of infarct core in acute ischemic stroke [56]. The infarct core is irrevocable while penumbra could be revocable. Quick and accurate segmentation of the penumbra is of great significance to the treatment of stroke. The lesion is usually defined by diffusion and perfusion MRIs [57]. SPES dataset contains 50 samples: 30 for training and 20 for testing. Each sample has multi-modality MRI sequences: T1c, T2, DWI, CBF, CBV, TTP and Tmax. All anonymous samples come from a hospital, and all sequences have been skull-stripped and rigidly registered. In SPES challenge, an automated method to segment and quantify the penumbra estimation is important for treatment. In our experiment, we use multi-modality MRIs (DWI,TTP,CBF,Tmax,T2 and CBV) as inputs.

Table 2
The ranking of the testing datasets on the SISS and SPES segmentation challenges. Values correspond to the mean (standard deviation).

	SISS Challenge				SPES Challenge			
Rank	Teams	DC[0,1]	ASSD(mm)	HD(mm)	Teams	DC[0,1]	ASSD(mm)	HD(mm)
1	lianl1	0.57(0.29)	8.22(16.25)	43.02(30.48)	maieo1	0.81(0.09)	1.36(0.74)	23.62(12.99)
2	clera1	0.56(0.29)	9.26(17.13)	33.64(27.85)	clera1	0.80(0.11)	1.43(1.04)	25.70(17.08)
3	fengc1	0.55(0.30)	8.13(15.15)	25.02(22.02)	lianl1	0.79(0.09)	1.79(0.54)	36.93(25.42)
4	martc2	0.50(0.32)	14.69(17.82)	80.06(22.00)	robbd1	0.78(0.09)	2.77(2.77)	40.27(25.10)
5	abdua1	0.43(0.31)	16.85(15.71)	74.66(25.10)	martc2	0.77(0.14)	1.78(0.84)	25.88(12.98)
6	robbd1	0.43(0.30)	14.22(14.41)	62.58(30.61)	dutif1	0.76(0.10)	2.24(0.79)	24.16(12.62)
7	maieo1	0.42(0.33)	17.59(21.06)	56.39(30.65)	fengc1	0.76(0.09)	2.29(1.76)	30.65(16.49)
8	haect1	0.37(0.33)	17.36(19.27)	63.59(31.68)	abdua1	0.72(0.24)	4.68(11.60)	32.76(20.66)
9	dutif1	0.35(0.31)	18.74(20.64)	55.99(35.09)	haect1	0.65(0.19)	3.87(3.05)	34.65(16.23)

4.3. Model training

In SISS and SPES challenges, we resize all MRI sequences to the scale of 160×160. In the image preprocessing stage, all 3D training images and testing images of two challenges are split into 2D images. The images of training datasets are augmented by flipped and randomly rotated methods. In the training process, the hyper-parameters are kept constant: the batch size is set to 8, the epoch is set to 60 and the learning rate is set to 0.001. In addition to these commonly used parameters, we added dropout parameters in convolutions. which is set to 0.1. In the testing process, network inherits the weight of the training model. After testing, all images are restored to the original size by the affine transform method. Then a post-processing step to refine the network output, we use the image median filtering algorithm [58] to alleviate noises and preserve the edge details of images. Finally, we synthesize the 2D slice images into 3D.

4.4. Result and analysis

We evaluate MK-DCNN on SISS and SPES challenges, and compare the results of 9 teams that participated in both challenges at the same time (at the moment of writing the paper). The results and rankings of the participants are on the challenge web page[1]. Our team name is lianl1. The results with the ranking of the on-site SISS and SPES are shown in Table 2. MK-DCNN has achieved very competitive rankings in both two segmentation tasks.

Among 9 submissions both on SISS and SPES challenges, MK-DCNN achieves superior performance on both challenges. Our team achieves top one ranking performance in SISS challenge, and we also achieves top 3 performance in SPES challenge. In SPES challenge, only team maieo1 and team clera1 are ahead of us. However, in SISS challenge, we surpassed team clera1 and far ahead of team maieo1.

4.5. Comparison with other segmentation methods

In addition, in SISS and SPES challenges, we also compare NK-DCNN with the state-of-the-art methods:U-net, FCN [45], EDD Net [59] and Res-FCN [32]. The comparison results are shown in Tables 3 and 4, respectively. Among the four comparison methods, the ranking of segmentation results is very unstable. For example, U-net and FCN methods performed well in the SPES challenge, but they did poorly in the SISS challenge. It is observed that MK-DCNN outperforms other four methods with DC, ASSD and HD metrics. Especially in the SPES challenge, compared with the second U-net method, MK-DCNN achieves an improvement of 0.4 in DC score, 1.55 in ASSD score and 27.18 in HD score. MK-DCNN is composed of U-net and dense blocks. Compared with U-net, dense blocks can help MK-DCNN extract more features. FCN combines semantic information from deep layers and shallow layers, which can produce accurate and detailed segmentation. MK-DCNN not only inherits the advantages of FCN, but also uses the multi-kernel strategy to obtain more receptive fields and contextual information, which can provide more information for correct segmentation. EDD Net is a deep fully convolutional network with mixed size image slices, which can help optimal lesion segmentation. Compared with EDD Net, MK-DCNN uses the dropout regularization method to alleviate the over-fitting problem and filter the number of feature vectors, which contributes to the performance gains. Compared with the residual block in Res-FCN, we use dense blocks to obtain more feature information from shallow layers. As shown in Tables 3 and 4, the results of all the methods in the SISS and SPES challenges are quite different. The results of U-net, FCN, EDD Net and Res-FCN in the SISS challenge are disappointing. These segmentation methods often do well in one specific medical images, do not generalize well to other types of medical images.

Table 3
The ranking of the testing datasets on the SISS segmentation challenge. Values correspond to the mean (standard deviation).

Methods	DC[0,1]	ASSD(mm)	HD(mm)
U-net[15]	0.14(0.16)	41.04(19.76)	83.00(21.00)
FCN[45]	0.17(0.24)	25.36(15.64)	82.09(19.03)
EDD Net[59]	0.49(0.33)	12.66(15.98)	59.12(26.85)
Res-FCN[32]	0.49(0.32)	12.41(16.41)	53.14(30.00)
MK-DCNN	0.57(0.29)	8.22(16.25)	43.02(30.48)

Table 4
The ranking of the testing datasets on the SPES segmentation challenge. Values correspond to the mean(standard deviation).

Methods	DC[0,1]	ASSD(mm)	HD(mm)
U-net[15]	0.75(0.12)	3.34(2.18)	64.11(29.14)
FCN[45]	0.70(0.12)	3.54(2.04)	60.42(28.56)
EDD Net[59]	0.68(0.16)	3.49(2.77)	51.86(24.93)
Res-FCN[32]	0.71(0.12)	3.43(1.94)	61.05(23.75)
MK-DCNN	0.79(0.09)	1.79(0.54)	36.93(25.42)

5. Ablationstudy

5.1. Data and pre-processing

In this section, to verify the superiority of MK-DCNN, we apply MK-DCNN to an acute stroke dataset with 29 stroke patients which come from an in-house MRI segmentation dataset. All scan sequences were collected by Phillips Achieve 3.0T MRI system. Each sample has multi-modality: T1, T2, DWI, Flair and ADC. The acquisition protocols of MRI sequences are shown as fol-

[1] http://isles-challenge.org.

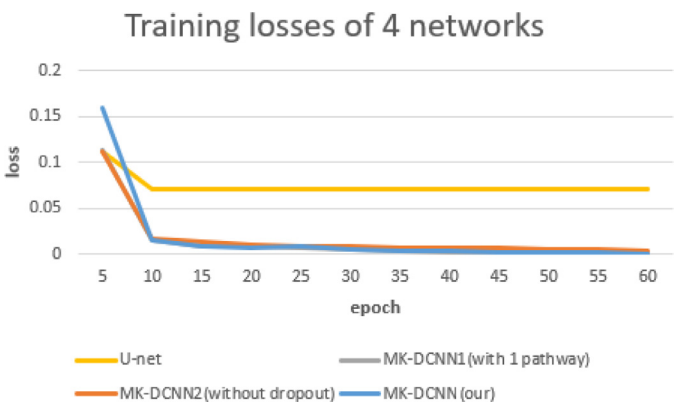


Fig. 7. Training losses of 4 networks.

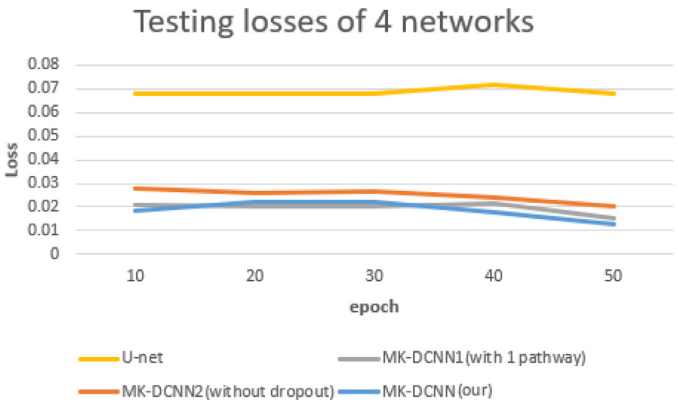


Fig. 8. Testing losses of 4 networks.

lows: matirx size (230 × 230 × 18), slices (18), slices spacing (1.0–1.5 mm), slices thickness (6 mm), echo time (87 ms), repetition time (23 ms). In image pre-processing, we use SPM12 software[2] to transform DICM images into NIfTI images. All sequences are co-registered to DWI sequence, and all processed MRI sizes are normalized to 160 × 160 pixels.

5.2. Ablation analysis

To evaluate the impact of dense block, multi-kernel and dropout regularization method in MK-DCNN, we compare the performance of U-net [15], MK-DCNN1 (with 1 pathway), MK-DCNN2 (two sub-networks and without dropout) and MK-DCNN (ours). In the infarct diagnosis of acute ischemic stroke, DWI is the most sensitive and common modality. We use DWI as input modality in ablation experiments. 22 samples are used as the training cases and 7 samples are used as the testing cases. Ablation experiments are carried out under the same parameter settings and the same data processing strategies for fair comparison.

As shown in Figs. 7 and 8, we analyze the training and testing leaning behaviors of 4 networks. It is clearly observed that the curves of MK-DCNN are lower than that of U-net. It means that MK-DCNN achieves a lower loss rate in both training and testing processes than that of U-net. We embed the dense blocks into U-net which can improve the performance of the network. As shown in Fig. 7, at the beginning of the training, MK-DCNN1 converges to a smaller loss rate than MK-DCNN does. From epoch 10 to 60, the loss rates of the two models almost coincide. However, in Fig. 8, the curve of the MK-DCNN shows the smaller loss rate than that of MK-DCNN1, which indicates that the multi-kernel has a greater contribution to the network than the single kernel. From Fig. 7, one can observe that the loss rate curves of MK-DCNN2 and MK-DCNN with very high coincidence on training dataset. However, in Fig. 8, it is obvious that the performance of MK-DCNN is better than MK-DCNN2, which validates the effectiveness of the dropout regularization method in MK-DCNN.

Table 5
Ablation study about performance of 4 networks with DC, ASSD and HD metrics.

Methods	DC[0,1]	ASSD(mm)	HD(mm)
U-net	0.53	4.32	3.01
MK-DCNN1	0.69	3.20	2.38
MK-DCNN2	0.62	4.17	2.57
MK-DCNN	0.74	2.01	2.38

In addition to analyze the loss rate of the 4 methods, we use DC, ASSD and HD as metrics to compare the performance on the testing dataset. Table 5 presents the segmentation results of 4 methods at the same configurations. It illustrates that MK-DCNN achieves better performance. Dense block, multi-kernel and dropout regularization method in MK-DCNN can effectively improve segmentation performance.

6. Discussion

In this study, we have proposed an automatic stroke MRI segment network MK-DCNN. MK-DCNN inherits the original advantages of the U-net and DenseNet121. The architecture of U-shape can improve the precise location and semantics capture of features. The dense block can reuse previous features, alleviate the vanishing-gradient problem and reduce the number of training parameters. In addition, two different kernels in the first layer are used to divide the network into two sub-networks to obtain different receptive fields, which helps MK-DCNN obtain much richer contextual information. Furthermore, we use the dropout regularization method to extract important information and reduce the number of neurons. We use four dropout rates (0.5, 0.2, 0.1 and 0) to validate the contribution of dropout regularization method in the proposed MK-DCNN. The experiments are tested on two public benchmark challenges: SISS and SPES. The results are shown in Table 6. It can be observed that the performance of MK-DCNN is the worst when the dropout rate is set to 0. Compared with dropout rates: 0.5, 0.2 and 0.1, the performance of MK-DCNN achieves the best in two challenges when dropout ratio is set to 0.1. Larger dropuout rate means more contextual information loss, which degrades the performance. Finally, we choose 0.1 as the value of dropout rate in proposed MK-DCNN.

In SISS and SPES challenges, although both sub-acute ischemic stroke and penumbra belong to stroke disease. However, they are essentially different in the prognosis. The infarct of stroke can not be saved, it obstructs blood supply and leads to tissue death. The tissue of penumbra surrounds the infarct, which is not completely obstructed. Penumbra could potentially be saved. It's important to prevent penumbra from transforming into infarct. In the clinical treatment decision, an automatic lesions segmentation method could provide reliable auxiliary evaluation for neurologists.

As shown in Tables 2–4, it is obvious that the DC, ASSD and HD scores are quite different in SISS and SPES challenges. Several top rank methods have achieved superior performance in SPES challenge. However, the results of the same methods in SISS challenge are quite different. The reason for this phenomenon is very complicated. Apart from the differences of disease and MRI sequences, according to our analysis, there are two main reasons leading to

[2] http://www.fil.ion.ucl.ac.uk/spm/.

Table 6
The results of different dropout rates used in MK-DCNN on the SISS and SPES segmentation challenges. Values correspond to the mean (standard deviation).

	SISS Challenge			SPES Challenge		
Dropout rate	DC[0,1]	ASSD(mm)	HD(mm)	DC[0,1]	ASSD(mm)	HD(mm)
0.5	0.52(0.32)	12.26(17.12)	50.34(30.30)	0.76(0.09)	2.10(0.57)	32.54(21.14)
0.2	0.56(0.30)	8.86(16.54)	38.41(29.42)	0.77(0.11)	2.20(0.78)	43.92(27.75)
0.1	0.57(0.29)	8.22(16.25)	43.02(30.48)	0.79(0.09)	1.79(0.54)	36.93(25.42)
0	0.51(0.34)	12.15(16.78)	51.37(30.30)	0.76(0.11)	2.78(1.34)	66.59(24.77)

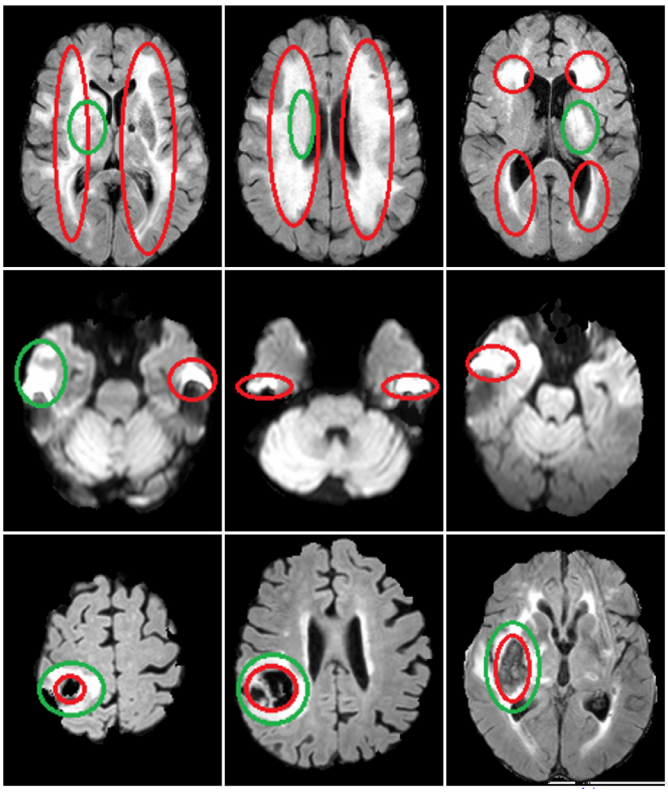


Fig. 9. Lesion examples in SISS. The first row shows three Flair slices with WMHs or demyelination lesions, red circles denote WMHs or demyelination and green circles denote stroke lesions respectively. The second row shows three DWI slices with artifacts, red and green circles denote artifacts and stroke lesions, respectively. The third row shows three Flair slices with hemorrhagic stroke and ischemic stroke, red and green circles denote hemorrhagic stroke and ischemic stroke lesions, respectively.

the low scores in SISS challenge. First, in SPES challenge, all of MRI sequences come from one medical center. MRI acquisition parameters are unified, and there is a slight difference in density and texture of the obtained MRIs. However, in SISS challenge, there are 64 samples, 28 training samples and 28 testing samples come from the same center, and 8 additional samples come from another center in the testing dataset. The 8 additional samples had different acquisition parameters and density, and there is no training data from the second center. Second, there are artifacts and signal consistencies of other lesions in MRI sequences of SISS, which makes big trouble for most methods. For example, infarcts of sub-acute stroke are strong signal in DWI and Flair modalities, and there are periventricular WMHs and demyelination lesions have isointense signal in the same MRI modalities, as shown in the first row in Fig. 9. The hyperintense artifacts also cause confusion in DWI modality, as shown in the second row in Fig. 9. The lesions of hemorrhagic stroke in the test samples also brought troubles to

Table 7
Multi-modality on SISS with three metrics. The bottom line is our final results.

Modalities	DC[0,1]	ASSD(mm)	HD(mm)
DWI+Flair+T1+T2	0.49(0.32)	12.41(16.41)	53.14(30.00)
DWI+Flair	0.57(0.29)	8.22(16.25)	43.02(30.48)

Table 8
Multi-modality on SPES with three metrics. The bottom line is our final results.

Modalities	DC[0,1]	ASSD(mm)	HD(mm)
DWI+CBF+CBV+T1c+T2+Tmax+TTP	0.76(0.10)	2.10(0.57)	32.54(21.14)
DWI+CBF+CBV+ T2+Tmax+TTP	0.79(0.09)	1.79(0.54)	36.93(25.42)

ischemic stroke segmentation. The pathologies of the two diseases are different, and the signals of two diseases are reversed in DWI modality, as shown in the third row in Fig. 9. However, there is no distinction between the two diseases in SISS challenge. All of these reflect the high variability of stroke characteristics provided by SISS challenge, which is a complex task for all participants. How to break the limitation of multi-data centers and lesion complexity problem is one direction of our future work.

In order to achieve better segmentation results, we use multi-modality in the SISS and SPES challenges. Multi-modality MRI sequences are commonly used in the medical image segmentation tasks, which can improve the accuracy of lesion segmentation [46,48,60]. In our experiments, we integrate the clinical experience of neurologists into the process of MRI modalities selection. In the diagnosis of sub-acute stroke, DWI and Flair are sensitive sequences, and the infarct areas show high signals on DWI, Flair and T2 sequences. However, the T2 sequences are fuzzy in SISS dataset and the signal of lesion is not obvious in T1. Finally, we choose DWI and Flair as input images in MK-DCNN in SISS challenge. In the clinical, the penumbra region is usually determined by PWI and DWI mismatch methods, although it is still controversial in medicine [61]. There is no PWI in SPES MRI sequences. In SPES challenge, combining medical prior knowledge, we use DWI, TTP, CBF, Tmax, T2 and CBV as input images in MK-DCNN. We used the different multi-modality sequences to conduct two comparative experiments on testing SISS and SPES datasets, respectively. The results are shown in Tables 7 and 8, which indicate that the effect of using all modalities as input data is not better than our selected multi-modality. The improper use of multi-modality may result in the increase of noise and decrease of the performance of the model. Better understanding of the pathological knowledge and incorporation with deep learning methods is a feasible strategy, it may better capture the heterogeneity feature and help promote segmentation performance.

7. Conclusion

In this study, we presented an end-to-end lesion segmentation network, and evaluated the performance of the network on two public benchmark challenges. In MK-DCNN, we elegantly com-

bined the U-shape with dense blocks, used multi-kernel and drop regularization method to achieve state-of-the-art performance on two challenges. Then we carried out the ablation experiments on an in-house dataset. The results demonstrate that the deep layer, multi-kernel strategy and dropout method in MK-DCNN can improve the performance of segmentation. The automatically generated segmentation of ischemic stroke lesion is important for clinical diagnosis. Our methods still need to be improved. There is a limitation in the segmentation of multi-data center and similar pathological tissues. This will be one direction of our future work. We would also try to analyze other stroke diseases by combining brain networks and hyper-graph techniques [62–66].

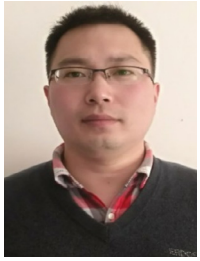
Acknowledgment

The work described in this paper was supported by the National Natural Science Foundation of China under Grant nos.61772557, 61772552, 61622213, 61728211; the 111 Project (No. B18059); the Hunan Provincial Science and Technology Program (2018WK4001)

References

[1] J. Grimaud, M. Lai, J. Thorpe, P. Adeleine, L. Wang, G.J. Barker, D.L. Plummer, P.S. Tofts, W.I. Mcdonald, D.H. Miller, Quantification of MRI lesion load in multiple sclerosis: a comparison of three computer-assisted techniques, Magn. Resonance Imaging 14 (5) (1996) 495–505.
[2] P. Schmidt, C. Gaser, M. Arsic, D. Buck, A. FÃ¶rschler, A. Berthele, M. Hoshi, R. Ilg, V.J. Schmid, C. Zimmer, An automated tool for detection of flair-hyperintense white-matter lesions in multiple sclerosis, Neuroimage 59 (4) (2012) 3774–3783.
[3] J. Mitra, P. Bourgeat, J. Fripp, S. Ghose, S. Rose, O. Salvado, A. Connelly, B. Campbell, S. Palmer, G. Sharma, Lesion segmentation from multimodal mri using random forest following ischemic stroke, Neuroimage 98 (9) (2014) 324–335.
[4] M.J. Cardoso, C.H. Sudre, M. Modat, S. Ourselin, Template-based multimodal joint generative model of brain data, Proceedings of the International Conference on Information Processing in Medical Imaging (IPMI), 2015, pp. 17–29.
[5] F. Forbes, S. Doyle, D. Garcia-Lorenzo, C. Barillot, M. Dojat, Adaptive weighted fusion of multiple MR sequences for brain lesion segmentation, Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI), 2010, pp. 69–72.
[6] M. Erihov, S. Alpert, P. Kisilev, S. Hashoul, A cross saliency approach to asymmetry-based tumor detection, Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015, pp. 636–643.
[7] J. Liu, Y. Pan, M. Li, Z. Chen, L. Tang, C. Lu, J. Wang, Applications of deep learning to MRI images: a survey, Big Data Min. Anal. 1 (1) (2018) 1–18.
[8] N. Lessmann, G.B. Van, M. Zreik, P.A. de Jong, B.D. de Vos, M.A. Viergever, I. Isgum, Automatic calcium scoring in low-dose chest CT using deep neural networks with dilated convolutions, IEEE Trans. Med. Imaging 37 (2) (2018) 615–625.
[9] L. Zhang, L. Lu, R.M. Summers, E. Kebebew, J. Yao, Convolutional invasion and expansion networks for tumor growth prediction, IEEE Trans. Med. Imaging 37 (2) (2018) 638–648.
[10] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: a deep convolutional encoder-decoder architecture for scene segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495.
[11] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, F.F. Li, ImageNet: a large-scale hierarchical image database, Proceedings of the IEEE Conference on Computer Vision and Pattern Reco (CVPR), 2009, pp. 248–255.
[12] K. Simonyan, A. Zisserman, in: Very deep convolutional networks for large-scale image recognition, 2014 arXiv preprint arXiv:1409.1556.
[13] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118.
[14] Y. Lécun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.
[15] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015, pp. 234–241.
[16] J.A. Golden, Deep learning algorithms for detection of lymph node metastases from breast cancer: helping artificial intelligence be seen, JAMA 318 (22) (2017) 2184–2186.
[17] H. Fu, J. Cheng, Y. Xu, D.W.K. Wong, J. Liu, X. Cao, Joint optic disc and cup segmentation based on multi-label deep network and polar transformation, IEEE Trans. Med. Imaging 37 (7) (2018) 1597–1605.
[18] O. Maier, B.H. Menze, d.G.J. Von, L. Hni, M.P. Heinrich, M. Liebrand, S. Winzeck, A. Basit, P. Bentley, L. Chen, ISLES 2015 - a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI, Med. Image Anal. 35 (2017) 250–269.
[19] J.A. Maldjian, J. Chalela, S.E. Kasner, D. Liebeskind, J.A. Detre, Automated ct segmentation and analysis for acute middle cerebral artery stroke, AJNR Am. J. Neuroradiol. 22 (2001) 1050–1055.
[20] A. Usinskas, R.A. Dobrovolskis, B.F. Tomandl, Ischemic stroke segmentation on ct images using joint features, Informatica 15 (2) (2003) 283–290.
[21] L.E. Poh, V. Gupta, A. Johnson, R. Kazmierski, W.L. Nowinski, Automatic segmentation of ventricular cerebrospinal fluid from ischemic stroke CT images, Neuroinformatics 10 (2) (2012) 159–172.
[22] C.R. Gillebert, W.G. Humphreys, D. Mantini, Automated delineation of stroke lesions using brain ct images, NeuroImage: Clin. 2 (2014) 540–548.
[23] N.M. Saad, S.A.R. Abubakar, S. Muda, M. Mokji, Fully automated region growing segmentation of brain lesion in diffusion-weighted MRI, IAENG Int. J. Comput. Sci. 39 (2) (2012) 674–677.
[24] Y. Kabir, M. Dojat, B. Scherrer, F. Forbes, C. Garbay, Multimodal MRI segmentation of ischemic stroke lesions, IEEE Eng. Med. Biol. Soc. (EMBS) 2007 (2007) 1595–1598.
[25] J. Mitraa, P. Bourgeata, J. Frippa, S. Ghosea, S. Rosea, O. Salvadoa, A. Connellyb, B. Campbellc, S. Palmerb, G. Sharmad, S. Christensend, L. Carey, Lesion segmentation from multimodal MRI using random forest following ischemic stroke, Neuroimage 98 (9) (2014) 324–335.
[26] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1–9.
[27] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS), 2012, pp. 1097–1105.
[28] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.
[29] C. Sun, S. Guo, H. Zhang, J. Li, M. Chen, S. Ma, L. Jin, X. Liu, X. Li, X. Qian, Automatic segmentation of liver tumors from multiphase contrast-enhanced ct images based on fcns, Artif. Intell. Med. 83 (11) (2017) 58–66.
[30] M. Drozdzal, G. Chartrand, E. Vorontsov, M. Shakeri, J.L. Di, A. Tang, A. Romero, Y. Bengio, C. Pal, S. Kadoury, Learning normalized inputs for iterative estimation in medical image segmentation, Med. Image Anal. 44 (2017) 1–13.
[31] K. Kamnitsas, C. Ledig, V.F.J. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Med. Image Anal. 36 (2016) 61–78.
[32] Z. Liu, C. Cao, S. Ding, T. Han, H. Wu, S. Liu, Towards clinical diagnosis: automated stroke lesion segmentation on multimodal mr image using convolutional neural network, IEEE Access 6, 2018, pp. 57006–57016.
[33] L. Liu, S. Chen, F. Zhang, F.-X. Wu, Y. Pan, J. Wang, Deep convolutional neural network for automatically segmenting acute ischemic stroke lesion in multi-modality, Neural Comput. Appl. (2019), doi:10.1007/s00521-019-04096-x.
[34] R. Guerrero, C. Qin, O. Oktay, C. Bowles, L. Chen, R. Joules, R. Wolz, M.C. Valdés-Hernández, D.A. Dickie, J. Wardlaw, White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks, Neuroimage Clin. 17 (2017) 918–934.
[35] R. Zhang, L. Zhao, W. Lou, J.M. Abrigo, V.C. Mok, W.C. Chu, D. Wang, L. Shi, Automatic segmentation of acute ischemic stroke from DWI using 3D fully convolutional densenets, IEEE Trans. Med. Imaging (2018), doi:10.1109/TMI.2018.2821244.
[36] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, Proceedings of the International Conference on Machine Learning (ICML), 2015, pp. 448–456.
[37] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (1) (2014) 1929–1958.
[38] G.E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, R.R. Salakhutdinov, in: Improving neural networks by preventing co-adaptation of feature detectors, 2012 arXiv preprint arXiv:1207.0580.
[39] X. Li, H. Chen, X. Qi, Q. Dou, C.W. Fu, P.A. Heng, H-denseunet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes, IEEE Trans. Med. Imaging (2018), doi:10.1109/TMI.2018.2845918.
[40] F. Milletari, N. Navab, S.A. Ahmadi, V-net: fully convolutional neural networks for volumetric medical image segmentation, Proceedings of the International Conference on 3d Vision (3D Vision), 2016, pp. 565–571.
[41] T. Brosch, L.Y.W. Tang, Y. Yoo, D.K.B. Li, A. Traboulsee, R. Tam, Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation, IEEE Trans. Med. Imaging 35 (5) (2016) 1229–1239.
[42] D.M. Pelt, J.A. Sethian, A mixed-scale dense convolutional neural network for image analysis, Proc. Natl. Acad. Sci. USA 115 (2) (2018) 254–259.
[43] G. Huang, Z. Liu, L.V.D. Maaten, K.Q. Weinberger, Densely connected convolutional networks, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2261–2269.
[44] X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS), 2011, pp. 315–323.
[45] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431–3440.

[46] X. Li, Q. Dou, H. Chen, C.W. Fu, X. Qi, D.L. Belav, G. Armbrecht, D. Felsenberg, G. Zheng, P.A. Heng, 3D multi-scale FCN with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi-modality mr images, Med. Image Anal. 45 (2018) 41–54.
[47] D. Nie, L. Wang, Y. Gao, D. Shen, Fully convolutional networks for multi-modality isointense infant brain image segmentation, Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI), 108, 2015, pp. 1342–1345.
[48] H. Chen, Q. Dou, L. Yu, J. Qin, P.A. Heng, Voxresnet: deep voxelwise residual networks for brain segmentation from 3D MR images, Neuroimage 170 (2017) 446–455.
[49] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. Davidson, S.P. Pereira, M.J. Clarkson, D.C. Barratt, Automatic multi-organ segmentation on abdominal ct with dense v-networks, IEEE Trans. Med. Imaging 37 (8) (2018) 1822–1834.
[50] W. Zaremba, I. Sutskever, O. Vinyals, in: Recurrent neural network regularization, 2014 arXiv preprint arXiv:1409.2329.
[51] X. Bouthillier, K. Konda, P. Vincent, R. Memisevic, in: Dropout as data augmentation, 2015 arXiv preprint arXiv:1506.08700.
[52] S.J. Nowlan, G.E. Hinton, Simplifying neural networks by soft weight-sharing, Neural Comput. 4 (4) (2014) 473–493.
[53] L.R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (3) (1945) 297–302.
[54] D. Kingma, J. Ba, in: Adam: a method for stochastic optimization, 2014 arXiv preprint arXiv:1412.6980.
[55] R.G. Gonzalez, J.A. Hirsch, W.J. Koroshetz, M.H. Lev, P. Schaefer, Acute ischemic stroke. imaging and intervention, J. Neuroradiol. 33 (3) (2006) 193.
[56] J.C. Baron, G. Marchal, Ischemic core and penumbra in human stroke, Stroke 30 (1) (1999) 93–99.
[57] G. Schlaug, A. Benfield, A.E. Baird, B. Siewert, K.O. LÃ¶vblad, R.A. Parker, R.R. Edelman, S. Warach, The ischemic penumbra: operationally defined by diffusion and perfusion MRI, Neurology 53 (7) (1999) 1528–1537.
[58] T. Huang, G. Yang, G. Tang, A fast two-dimensional median filtering algorithm, IEEE Trans. Acoust. Speech Signal Process. 27 (1) (1979) 13–18.
[59] C. Liang, P. Bentley, D. Rueckert, Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks, Neuroimage Clin. 15 (2017) 633–643.
[60] S. Pereira, R. Meier, R. Mckinley, R. Wiest, V. Alves, C.A. Silva, M. Reyes, Enhancing interpretability of automatically extracted machine learning features: application to a RBM-random forest system on brain lesion segmentation, Med. Image Anal. 44 (2018) 228–244.
[61] M. Goyal, B.K. Menon, C.P. Derdeyn, Perfusion imaging in acute ischemic stroke: let us improve the science before changing clinical practice, Radiology 266 (1) (2013) 16–21.
[62] X. Zhu, S. Zhang, R. Hu, Y. Zhu, J. Song, Local and global structure preservation for robust unsupervised spectral feature selection, IEEE Trans. Knowl. Data Eng. 30 (3) (2017) 517–529.
[63] J. Liu, M. Li, W. Lan, F.-X. Wu, Y. Pan, J. Wang, Classification of Alzheimer's disease using whole brain hierarchical network, IEEE/ACM Trans. Comput. Biol. Bioinf. 15 (2) (2018) 624–632.
[64] Y. Zhu, X. Zhu, M. Kim, J. Yan, D. Kaufer, G. Wu, Dynamic hyper-graph inference framework for computer assisted diagnosis of neurodegenerative diseases, IEEE Trans. Med. Imaging (2018), doi:10.1109/TMI.2018.2868086.
[65] Y. Kong, J. Gao, Y. Xu, Y. Pan, J. Wang, J. Liu, Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier, Neurocomputing 324 (9) (2019) 63–68.
[66] J. Liu, X. Wang, X. Zhang, Y. Pan, X. Wang, J. Wang, MMM: classification of schizophrenia using multi-modality multi-atlas feature representation and multi-kernel learning, Multimed. Tools Appl. 77 (22) (2018) 29651–29667.



Liangliang Liu received his M.S. degree from Henan University in 2014. He is currently a Ph.D. candidate in School of Computer Science and Engineering, Central South University, Changsha, Hunan, P.R. China. His research interests include machine learning, deep learning and medical image analysis.

Fang-Xiang Wu received the B.Sc. degree and the M.Sc. degree in applied mathematics, both from Dalian University of Technology, Dalian, China, in 1990 and 1993, respectively, the first Ph.D. degree in control theory and its applications from Northwestern Polytechnical University, Xi'an, China, in 1998, and the second Ph.D. degree in biomedical engineering from University of Saskatchewan (U of S), Saskatoon, Canada, in 2004. During 2004–2005, he worked as a Postdoctoral Fellow in the Laval University Medical Research Center (CHUL), Quebec City, Canada. He is currently a Professor of the Division of Biomedical Engineering and the Department of Mechanical Engineering at the U of S. His current research interests include computational and systems biology, genomic and proteomic data analysis, biological system identification and parameter estimation, applications of control theory to biological systems. He is serving as the editorial board member of five international journals, the guest editor of several international journals, and as the program committee chair or member of several international conferences. He has also reviewed papers for many international journals.

Jianxin Wang received his B.S. and M.S. degree in Computer Science from Central South University of Technology, P. R. China, and his Ph.D. degree in Computer Science from Central South University. Currently, he is the vice dean and a professor in School of Computer Science and Engineering, Central South University, Changsha, Hunan, P.R. China. He is currently serving as the executive editor of International Journal of Bioinformatics Research and Applications and serving in the editorial board of International Journal of Data Mining and Bioinformatics. He has also served as the program committee member for many international conferences. His current research interests include algorithm analysis and optimization, parameterized algorithm, bioinformatics and computer network. He has published more than 200 papers in various International journals and refereed conferences. He is a senior member of the IEEE.