

Research Article

DMCNN: A Deep Multiscale Convolutional Neural Network Model for Medical Image Segmentation

Lin Teng ¹, Hang Li ¹ and Shahid Karim ²

¹Software College, Shenyang Normal University, Shenyang 110034, China

²Institute of Image and Information Technology, Harbin Institute of Technology, Harbin 150000, China

Correspondence should be addressed to Hang Li; lihangsoft@163.com and Shahid Karim; shahidhit@yahoo.com

Received 29 September 2019; Revised 21 November 2019; Accepted 28 November 2019

Guest Editor: Liang Zhao

Copyright © 2019 Lin Teng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medical image segmentation is one of the hot issues in the related area of image processing. Precise segmentation for medical images is a vital guarantee for follow-up treatment. At present, however, low gray contrast and blurred tissue boundaries are common in medical images, and the segmentation accuracy of medical images cannot be effectively improved. Especially, deep learning methods need more training samples, which lead to time-consuming process. Therefore, we propose a novelty model for medical image segmentation based on deep multiscale convolutional neural network (CNN) in this article. First, we extract the region of interest from the raw medical images. Then, data augmentation is operated to acquire more training datasets. Our proposed method contains three models: encoder, U-net, and decoder. Encoder is mainly responsible for feature extraction of 2D image slice. The U-net cascades the features of each block of the encoder with those obtained by deconvolution in the decoder under different scales. The decoding is mainly responsible for the upsampling of the feature graph after feature extraction of each group. Simulation results show that the new method can boost the segmentation accuracy. And, it has strong robustness compared with other segmentation methods.

1. Introduction

Medical imaging makes a critical difference in clinical diagnosis [1–3]. Recently, with the progress of medical imaging technology and the continuous development of artificial intelligence image processing, medical image processing technology has gradually developed into a key research field. It is vital in clinical application. The aim of medical image segmentation technology is to segment the interested part by some deep automatic segmentation algorithms and make the segmentation results as close as possible to the original structure of the region [4]. Segmentation of medical image has big significance in clinical diagnosis and pathological diagnosis. Measuring lesion volume with segmented images can assist doctors to determine the disease and make treatment plans [5].

Medical imaging segmentation (MIS) is an indispensable stage in ROI (region of interest) extraction, quantitative analysis, and 2D reconstruction. The images will be segmented with the same or similar features (such as intensity,

color, and texture) into separated areas, particularly to extract the lesion areas with special meanings or other regions of interest (ROI) from the complex background, so as to provide basis for clinical analysis [6]. Magnetic Resonance Imaging (MRI) uses the principle of nuclear magnetic resonance, which not only has high soft tissue resolution but also provides rich and high-resolution three-dimensional brain tissue information. Therefore, how to segment the medical images accurately in MRI images is becoming a challenging task in medical image research [7].

Through the analysis of research status, we summarize three kinds of traditional MIS methods: (1) manual segmentation method, which is tedious, excessive labor, subjective, prone to error, and not suitable for large-scale research [8]; (2) semiautomatic segmentation method, which requires accurate control of prior parameters and consumes much time in the process of parameter tuning [9]; (3) the traditional segmentation methods such as graph-based, deformation model, and active appearance model [10–12], which is based on simple registration method.

However, due to the differences between the hippocampus, in terms of the segmentation efficiency and accuracy, a simple registration method is still not ideal.

Currently, deep learning has attracted more attention, and the model based on deep CNN and its variants have been diffusely used in various fields of medical image processing and also achieved better results [13–15]. For example, Cha [16] presented MR brain image automatic segmentation method using a CNN network. The method was independent on explicit features only requiring a single MR (magnetic resonance) image. Lu et al. [17] used two-dimensional convolutional neural network to evaluate the segmentation of electron microscope images. Zhang et al. [18] adopted deep CNN to evaluate the segmentation of multimodal brain images. Although the models based on CNN have obtained better performance, these methods have a common problem, namely, all networks take image block as input, due to a large amount of overlapping image blocks, the redundant computation will increase the time cost for testing the network, and the image block size will influence the capability of the trained network. To solve the problem of image segmentation, many researchers have come up with many approaches based on the fully convolutional network (FCN) model to remedy limitations of image segmentation. FCN can take the entire image as the input of the network and generate the corresponding output of the whole image, thus avoiding the problems caused by the use of image blocks. However, it had low efficiency. So, a deep multiscale CNN model is put forward for medical image segmentation. The major contributions are illustrated as follows:

- (1) First, we extract the regions of interest from the raw medical images. Then, the data augmentation is operated to obtain more training dataset. The encoder, U-net, and decoder models are used for constructing our proposed segment framework.
- (2) Encoder is mainly responsible for feature extraction of 2D image slice.
- (3) In different scales of the decoder, it will acquire the features of each encoder's block by deconvolution operation, and then the U-net joins them together.
- (4) The decoding is mainly responsible for the upsampling of the feature graph after feature extraction of each group.

This paper is originated as follows. In Section 2, related works are introduced for the image segmentation. Section 3 describes the proposed DMCNN in detail. In Section 4, we conduct experiments and give analysis. Section 5 concludes the work.

2. Related Works

2.1. Inception Model. In order to make the convolutional neural network have better learning ability, the most direct and effective method is to make the network layer deeper. However, there are some disadvantages in this operation: (1) if the training number and dataset are limited, more

parameters will easily lead to overfitting; (2) if the network is larger, it is hard to utilize due to the greater endless computation; (3) the deeper the network is, the gradient will disappear, which leads to the diffusion of gradient. Under this situation, it is difficult to optimize the network model. Inception v1 is proposed in 2014 [19], the convolution layer of 1×1 , 3×3 , and 5×5 , and pooling layer of 3×3 are stacked together, which increases the width of the network and also enhances the adaptability of the network in terms of the scale. This operation can extract features from different scales. An important improvement in Inception v3 is decomposition, the two-dimensional convolution of $N \times N$ is divided into two-dimensional convolutions of $1 \times N$ and $N \times 1$. The advantage of this method is that it can not only accelerate the computation but also increase the nonlinearity of the network.

2.2. Batch Normalization. BN aims to add a standardized processing for the input data of each layer in the training process of neural network, which also belongs to the network layer. Previously, we mentioned that in addition to the output layer of the network, the parameters of the lower layer of the network are updated during the training, which caused the change of the distribution of the input data in the latter layer. In each layer, it is better to add a preprocessing operation. For example, the data in the third layer of network are normalized. Then, it inputs the third layer for calculation, so that we can solve the problem of "Internal Covariate Shift" [20–22]. By introducing batch standardization method, the network's convergence speed will be greatly increased. The overfitting can also be controlled. The dropout and regularization operation will be realized with little utilization.

2.3. End-To-End Models and Jump Connections. Compared to the traditional image block-based convolutional neural network model, the end-to-end model utilizes the entire image as input. The entire image will be generated as output [23, 24]. The image block-based model needs to foretell each size of pixel in the slice separately. Therefore, the end-to-end model adopted in this paper can evidently decrease the time consumption when segmenting images. Generally, end-to-end models primarily include fully CNN and faster CNN. It combines the feature maps of different levels. Unlike FCN, the U-net model adopts jump connection to join the feature information obtained from the shrinking coding path and the deconvolution operation in the expanded path together which is favourable for obtaining multiscale feature information to strengthen the network's feature extraction ability.

3. Proposed DMCNN

3.1. Data Preprocessing. For the acquired medical images, the region of interest should be extracted and preprocessed to serve as the samples for training and testing the network. After that, gray level regularization is carried out on all acquired ROIs. And, the mean and SD (standard deviation)

are measured. The gray level regularization is assessed by subtracting mean value and dividing by SD. In the subsequent training and testing process, the extracted images are transported into the network model as samples, and the extracted region is shown in Figure 1. For example, the size of the raw image is 256×256 , and after ROI extraction, we get the 128×128 patch and input it into network for training.

3.2. Data Extension. Since labeled public medical image datasets are little online and they are inconvenient to use, training a deep CNN model is troublesome [25]. In our proposed model, we first adopt some newest data argumentation methods to expand the raw data in order to increase the number of available training samples. In this paper, five data expansion methods are adopted including vertical direction reversal, random angle rotation, random translation, horizontal direction reversal, and image deformation. Figure 2 shows an example of the image expansion.

3.3. Proposed Network Model. To raise the accuracy of medical image segmentation, spatial information of images and relevant information between 2D slices are effectively utilized. The proposed deep multiscale convolutional neural network model contains three parts: encoder, U-net, and decoder as shown in Figure 3.

3.3.1. Encoder. It is mainly responsible for feature extraction of 2D slices, whose network structure is shown in Figure 4. Small convolution kernel in convolutional network is conducive to capturing local information, while large convolution kernel is conducive to capturing global information. However, ROIs are different in 2D slices, and it is difficult to select an accurate and universal convolution kernel. For this purpose, we use three different convolution layers (1×1 , 3×3 , 5×5) to extract information of multiple scales in Inception V1, which can extract more features. Additionally, to reduce the amount of computation, asymmetric convolution kernel is used in this experiment to decompose the $N \times N$ two-dimensional convolution into two one-dimensional convolutions with $1 \times N$ and $N \times 1$.

Meanwhile, to expand the receptive field of convolution and perfectly obtain multiscale information without increasing the size of parameters, this paper adds dilated convolutions with expansion coefficients of 2 and 4, respectively. For an ordinary convolution layer of 3×3 , the receptive field of its convolution kernel is 3×3 . In this paper, after the employment of dilated convolution (DC), the size of parameters remains unchanged, but the receptive field of the convolution kernel becomes 7×7 and 15×15 . It can be seen that DC greatly increases the receptive field of the convolution layer without increasing parameters number. As shown in Figure 4, there are three kinds of dilated convolution with expansion coefficients of 1, 2, and 4, respectively, in the encoded part, and the receptive fields of the corresponding convolution kernel are 3×3 , 7×7 and 15×15 , respectively. After using the three receptive fields

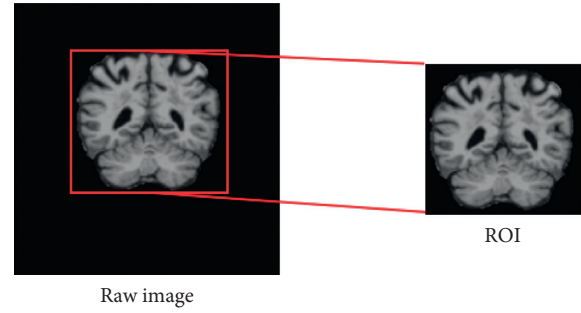


FIGURE 1: Extracted ROI.

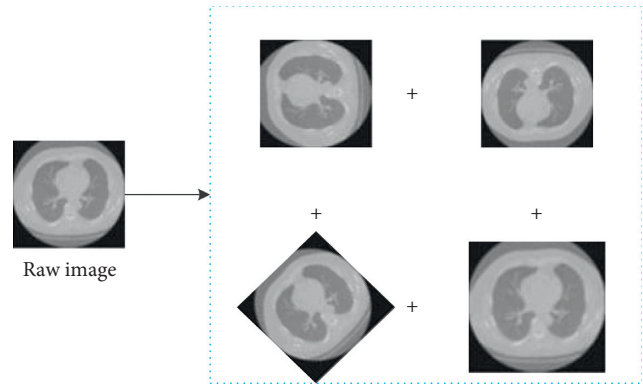


FIGURE 2: Proposed medical image segment model.

with different sizes, it is not only conducive to feature extraction but also conducive to better capturing the pathological area features.

Then, after cascading the feature graphs extracted from five different convolution layers, we utilize two ordinary 3×3 convolution layers in feature extraction process. Finally, to reduce the size of the feature map, it connects a maximum pooling layer. In Figure 4, the number of channels in each convolution layer is 16. Considering the convergence of networks, batch normalization is added behind the convolution layer, and ReLu layer is used as the activation function.

3.3.2. U-Net Model. Figure 5 shows the proposed U-net network structure employed. The U-net network structure contains two parts as follows:

- (a) Contractile encoder part on the left. It processes input medical images.
- (b) Decoder part on the right. It produces labeled output.
- (c) The skip connection. It can cascade the features of each block in the encoder. Features are obtained by deconvolution operation in the decoder.

The U-net model in this paper is displayed as Figure 5.

The proposed entire U-net network consists of twenty-eight convolution layers. In here, twenty-four convolution layers are spread over four convolution blocks and four deconvolution blocks. The contraction encoder in deep

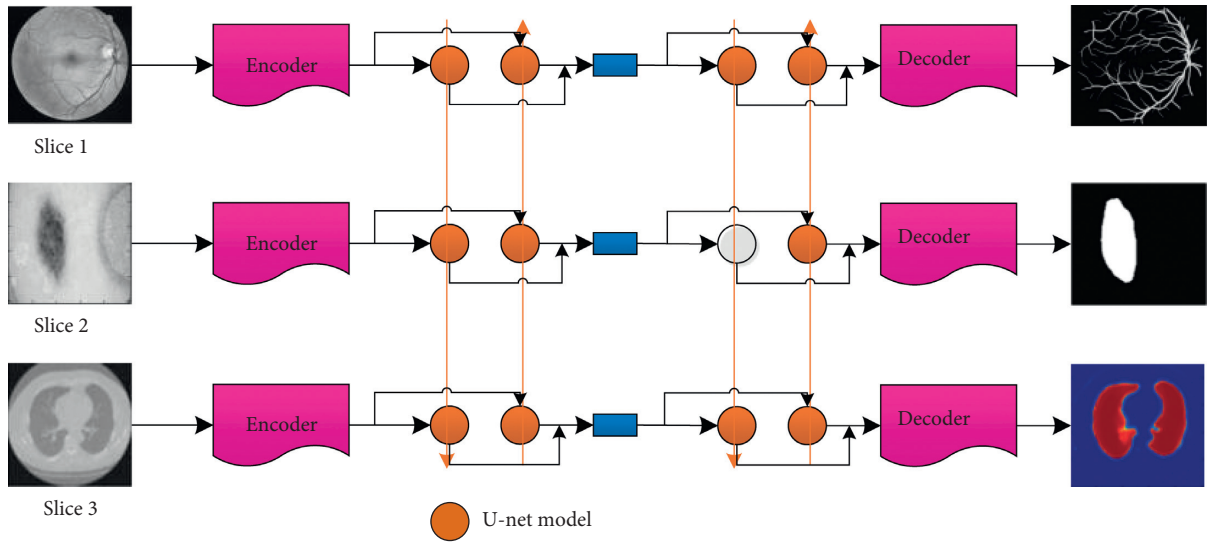


FIGURE 3: Proposed medical image segment model.

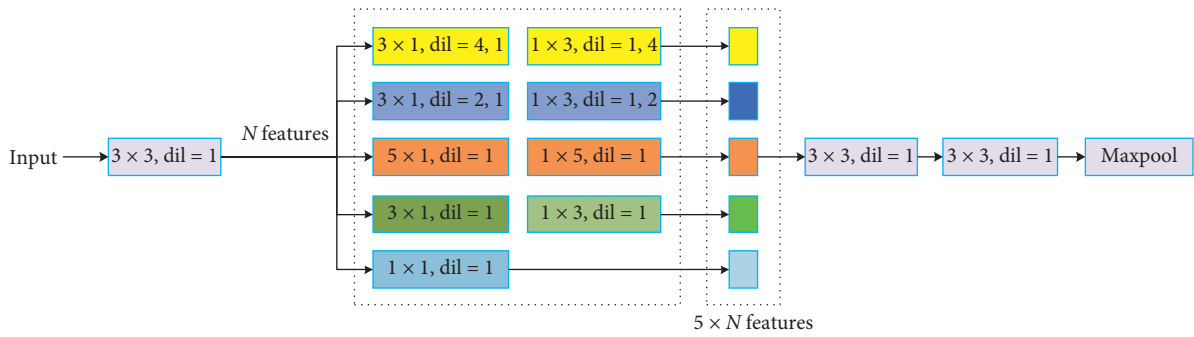


FIGURE 4: Encoder structure.

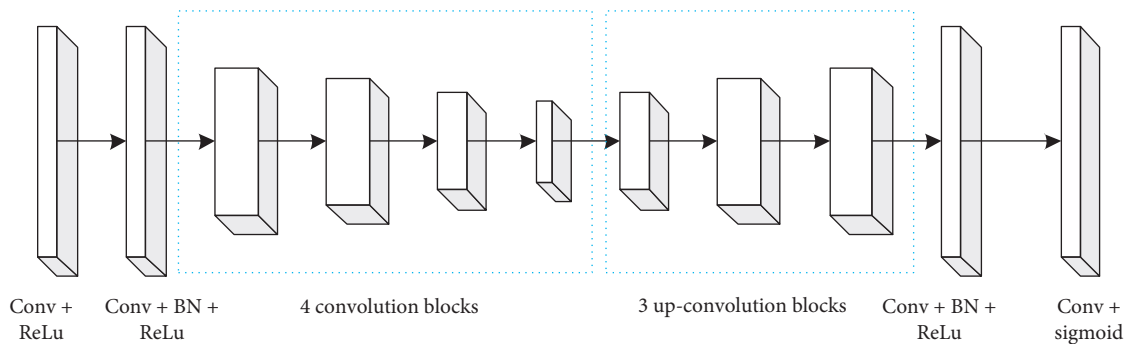


FIGURE 5: Deep multiscale U-net model.

multiscale U-net covers four convolution blocks shown in left of Figure 4. Each convolutional block contains two convolutional layers (conv). Each convolutional layer utilizes a 3×3 convolutional kernel to carry out the convolution. The step size is 1. Synchronously, each convolutional layer follows a BN layer and a ReLu layer to modify the network performance. ReLu activation function has sparse ability, and it can better learn the relatively sparse features from the effective data dimension and play the role of feature

automatic decoupling. In each convolution block, its first convolution layer can double feature graphs. The number of feature graphs will be increased to 64. After four convolution layers, the number of feature graphs is increased from 64 to 1024. Between each convolution block, the sampling method of traditional U-net adopts maximum pooling. In this paper, we use the 2×2 convolution kernel with step length 2 for down-conv operation to achieve a convolution block feature on the image of the sampling operation. Through this down-

conv operation, the size of feature graph is reduced by half with iterative deepening, so that the size of input original image is decreased from 128×128 layer by layer to 8×8 .

The expansion decoder of the U-net model contains three deconvolution blocks as shown in the right of Figure 6. The deconvolution up-conv operation adopts the 3×3 convolution kernel. The step size is 2, and the size of the feature graph is increased by twice the original size through the deconvolution operation. This process can recover the feature graph as the raw input image in the last deconvolution block. Meanwhile, the number of feature graphs is halved after each deconvolution operation. The feature graphs obtained by deconvolution are cascaded with the corresponding feature graphs in the convolution block as the feature input of the deconvolution block. Two convolution layers are in each deconvolution block. The 3×3 convolution kernel is utilized (the step size is 1). The first convolution layer will reduce the number of feature graphs by half after every cascading. Not exactly the same as the original U-net structure, the presented U-net structure is filled with zero filler in each convolution layer. According to formula (1), the output size of the deep multiscale CNN model can be guaranteed to be consistent with the input image data size by using zero padding:

$$I_{\text{output}} = \frac{(I_{\text{input}} - F + 2P)}{S + 1}, \quad (1)$$

where I_{input} and I_{output} are on behalf of the input and output images' size in DMCNN, respectively. F represents the convolution kernel with size 3×3 . P denotes the fill size with 1×1 . $S = 1$ stands for step size in this paper.

At the end of the proposed U-net model in this paper, we skillfully adopt a convolution layer (whose size is 1×1) to lessen the number of feature graphs to 1. The final output will be disposed by the Sigmoid function. We can obtain the value of each pixel between 0 and 1. The lesion area is a probability distribution. Through the above processing, the final image is considered as the probability graph of DMCNN. The value corresponding to each pixel indicates the probability that the point belongs to the lesion.

3.3.3. Decoding Part. It is mainly responsible for upsampling the feature graph after extracting each group feature, and the structure is shown in Figure 7. The decoder section contains one deconvolution layer and one convolution layer. Both deconvolution and convolution have batch-normalization and ReLu. After the upsampling, the feature graph becomes the same resolution as the input image. Finally, the final segmentation result is obtained by softmax classifier to analyze the end-to-end segmentation.

3.4. Loss Function. Different from the commonly used pixel point-based softmax loss function [26, 27], Dice loss function is based on region loss function. In medical image segmentation, Dice index is often used to measure the overlap rate between the object and the detection area. If the Dice value is larger, then the overlap degree is higher, and

the segmentation effect is better. However, Dice index cannot be directly used as a loss function, so we use the improved Dice function. Dice function is a function that gives feedback to network parameters after independent evaluation for each area [28]. The calculation form and process of Dice function are in good agreement with medical image segmentation. Therefore, Dice loss function used in this paper is defined as follows:

$$\text{Dice}(g, p) = 1 - \frac{2\sum_i^v p_i g_i}{\sum_i^v (p_i^2 + g_i^2)}, \quad (2)$$

where g stands for the ground truth. p is the predicted value. v is the number of pixels in each image block. Dice always is used as a loss function, when comparing the probability graph with the labeled. The background part whose labeled value is 0 will not be calculated into the loss to avoid the situation of unbalanced category and accelerate the convergence of the network and improve the segmentation accuracy.

4. Results and Discussion

4.1. Dataset and Evaluation Index. The dataset is from ADNI (Alzheimer's Disease Neuroimaging Initiative: adni.loni.usc.edu) [29, 30]. In this advanced researches, 100 groups of brain MRI images and segmented hippocampal tags are obtained from ADNI library. From this group, 80 groups are randomly selected for cross-validation, and the remaining 20 groups are for testing.

To improve the segmentation veracity, this study pre-processes the data with three steps. First, consider that the hippocampus only accounts for a small part of the whole brain MRI image and other parts are invalid areas. The pixel values in brain MRI are statistically analyzed, and the images are cropped into $80 \times 80 \times 40$ including the hippocampus and the blank area around it. In this way, invalid background information can be reduced without any influence on the integrity of valid information. Second, to accelerate the convergence of the network and consider the inconsistency of the pixel values of MRI images in ADNI, the mean and SD methods are utilized to normalize the images. Thirdly, making allowances for the small number of samples in the dataset, we enhance the obtained MRI images by left rotation and right rotation and finally obtain 400 MRI images.

To accurately reflect the performance differences between algorithms, we use uniform platform. The hardware environment of the experiment is NVIDIA GTX1060Ti, Intel Corei7 processor, and the software environment was Keras2.2.4. In the experiment, glorot normal distribution method is used to initialize the weight. The image size is 300×300 pixel used in this section. Execution environment is GPU and Geforce GTX 1060. The parameters used in DMCNN are given in Table 1.

To quantitatively evaluate the performance of the new approach, dice similarity coefficient (DSC), sensitivity (SEN), and predictive positivity value (PPV) are selected as the evaluation indexes for the medical image segmentation result. They are defined as follows:

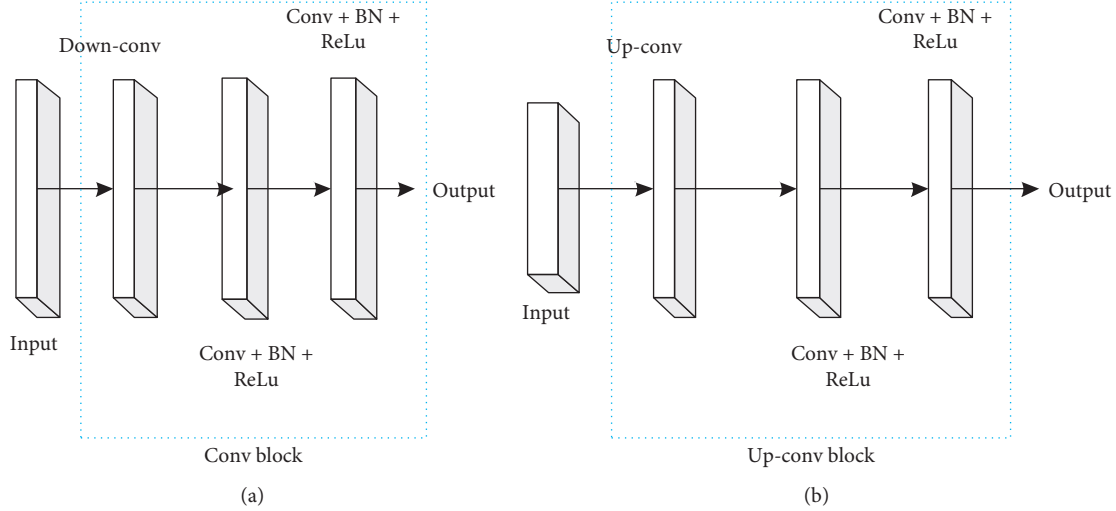


FIGURE 6: Convolution block and Up-convolution block structure.

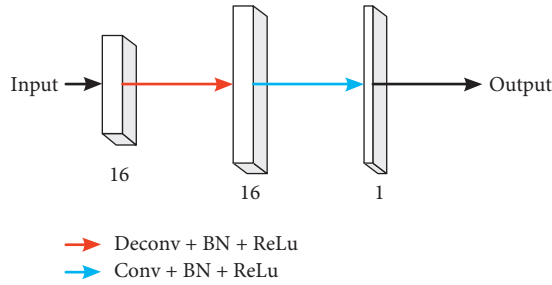


FIGURE 7: Decoder structure.

TABLE 1: Parameters in this experiment.

Learning rate	0.001
Batch size	8
L2 regularization	0.0001

TABLE 2: *Hippocampus* segment results with DMCNN.

Images	DSC	SEN	PPV
100	88.37	89.65	89.54
400	90.58	91.92	92.73

$$\begin{aligned}
 \text{DSC} &= 2 \frac{P \cap T}{P + T}, \\
 \text{SEN} &= \frac{P \cap T}{T}, \\
 \text{PPV} &= \frac{P \cap T}{P},
 \end{aligned} \tag{3}$$

where P denotes the lesion region segmented by the presented algorithm. T expresses the region of Ground truth. $P \cap T$ represents the pixel region of the intersection between the algorithm's segmentation region and the true segmentation region.

TABLE 3: *Hippocampus* segment results with DMCNN and single-scale CNN.

Method	DSC	SEN	PPV
Single-scale CNN	86.54	86.95	87.31
Multiscale CNN	91.26	90.89	91.57

TABLE 4: *Hippocampus* segment results with different U-net models.

Method	DSC	SEN	PPV
U-net	89.26	88.73	89.45
2D U-net	89.65	89.21	90.14
Proposed	91.23	90.87	91.58

TABLE 5: *Hippocampus* segment results with different methods.

Method	DSC	SEN	PPV	Time
TLWK	84.62	83.17	86.54	11.5 s
MNF	87.13	86.45	87.53	11.8 s
SUSAN	88.62	88.92	89.85	10.2 s
DMCNN	92.54	91.87	92.08	8.5 s

4.2. Comparative Analysis of Segmentation. In this paper, 100 images before augmentation and 400 images after augmentation are segmented by the proposed method. The evaluation indexes in above section are used for evaluation. The comparison results of DSC, SEN, and PPV are given in Table 2.

From Table 2, it reveals that data augmentation can greatly improve the segmentation accuracy. This also fully proves the importance of datasets in the deep learning model construction process. The size of the datasets can directly affect the learning capability of the model.

To verify that DMCNN can effectively capture information between slices, we conduct the comparison between multiscale convolutional neural network and single-scale CNN in this paper. Other conditions remain unchanged.

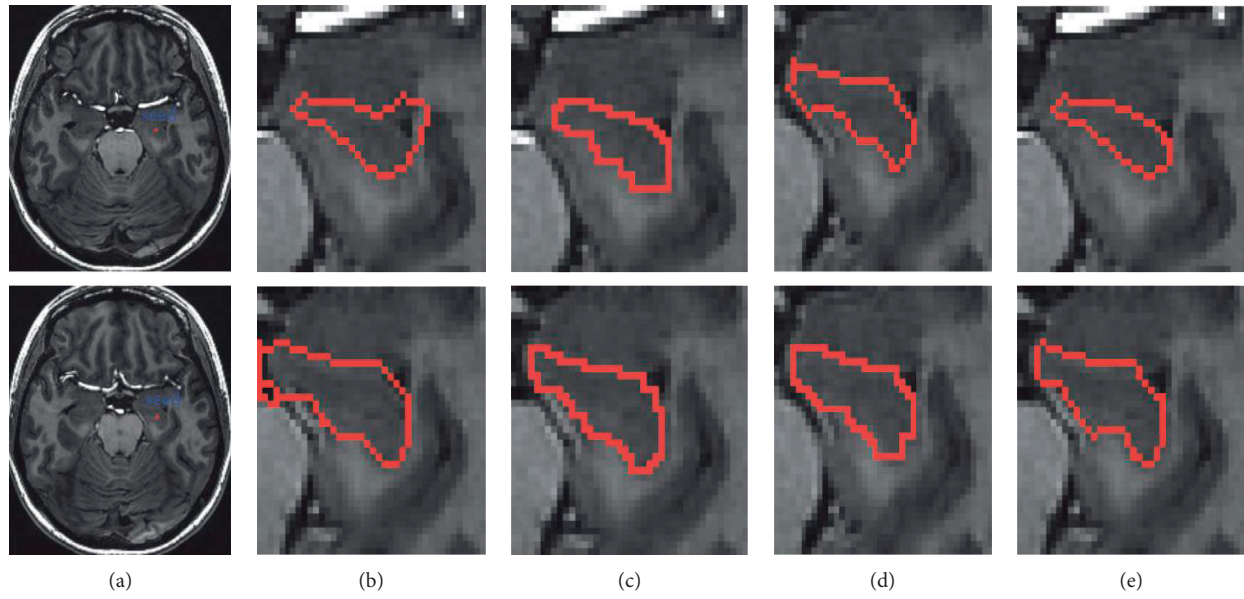


FIGURE 8: Segmentation of hippocampus with different algorithms. (a) Original images of hippocampus with initial seed points (blue). Results of the (b) TLWK model; (c) MNF method; (d) SUSAN method; (e) DMCNN method.

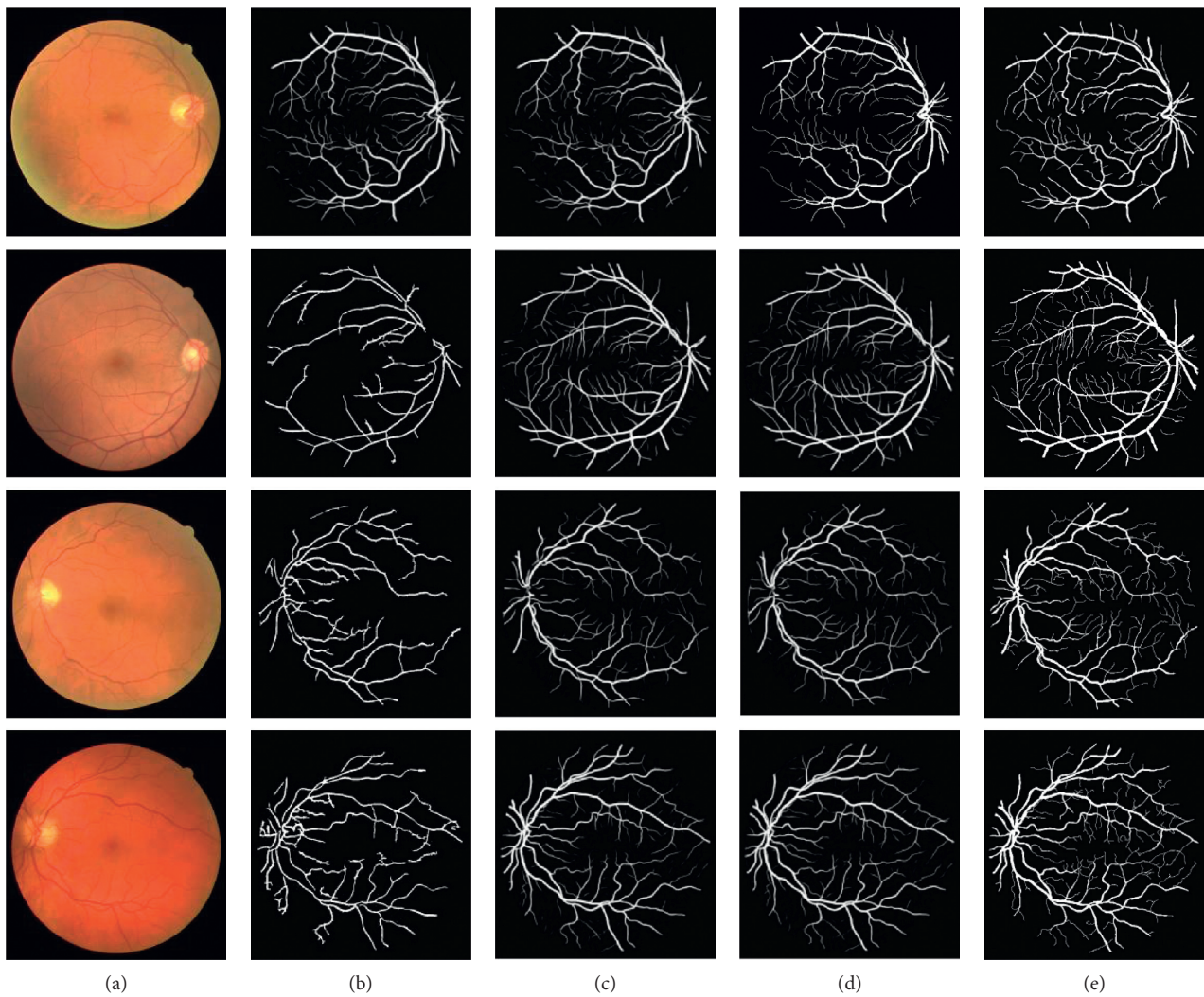


FIGURE 9: Segmentation results: (a) original image; (b) TLWK method; (c) MNF method; (d) SUSAN method; (e) DMCNN method.

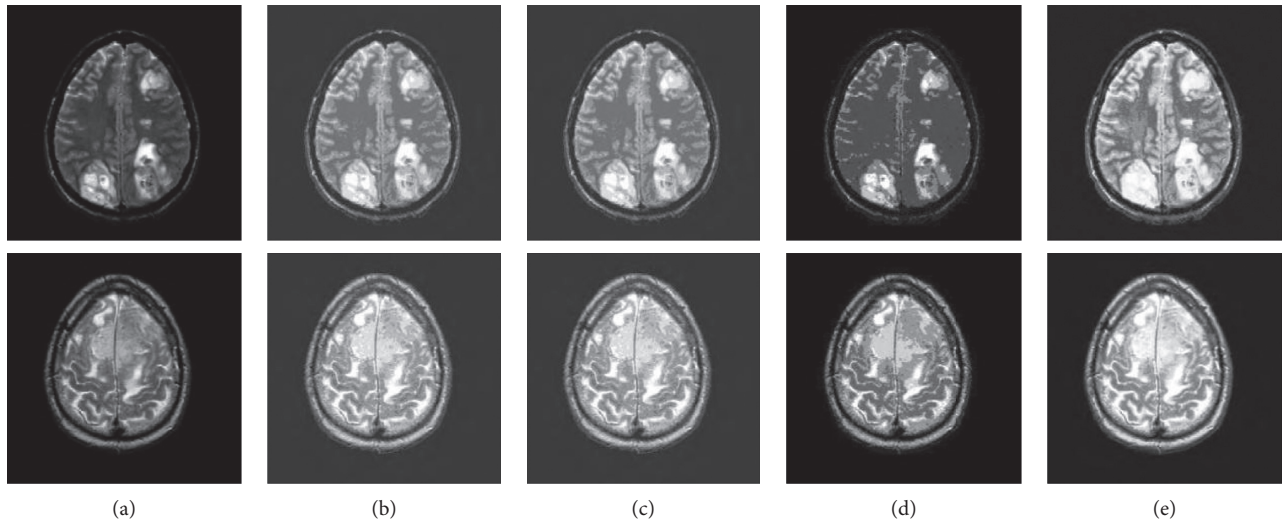


FIGURE 10: Segmentation results: (a) original image; (b) TLWK method; (c) MNF method; (d) SUSAN method; (e) DMCNN method. The first row is sarcoma, and the second is meningioma.

TABLE 6: IoU results with different methods.

Method	<i>Hippocampus</i>	Retina	Sarcoma	Meningioma
TLWK	72.48	81.73	75.48	73.69
MNF	78.31	84.56	79.35	78.24
SUSAN	82.37	88.57	82.54	81.53
DMCNN	89.62	90.24	89.66	88.74

The effectiveness of the proposed model can be observed from Table 3.

We can see that the segmentation accuracy of DMCNN is significantly higher than that of single-scale CNN, which further verifies that DMCNN can better learn more feature information between slice sequences than single-scale CNN.

Meanwhile, we study the effect of different network models on experiment results. Two representative segmentation methods including U-net and 2D U-net network model are compared with our deep multiscale U-net given in Table 4. The segmentation accuracy obtained by the DMCNN method is higher than that of the other two methods, indicating that it can extract features more efficiently and improve the segmentation accuracy.

Compared with the multiple groups of up- and down-sampling layers in U-net and 2D U-net networks, the proposed network model only contains one up- and down-sampling layer, which greatly reduces the size of the parameters. The number of parameters of the encoding part and decoding part is below 5000, which significantly reduces the computation time in this article.

We also conduct comparison experiment with state-of-the-art segment methods including TLWK [31], MNF [32], and SUSAN [33] on our medical data. The results are given in Table 5.

TLWK adopted the traditional random forest regression method, and MNF adopted the multiscale method. They are all automatic segmentation methods, due to the large gap between different individuals, and the accuracy and

efficiency of segmentation are often not ideal, which is not as high as the precision of the automatic segmentation method in this paper. SUSAN simply improves 2D U-net, so the results are not very good. In general, the proposed method combining convolution neural network and multiscale U-net model is superior to other current methods for medical image segmentation. And, the time consumption is shorter than other methods too.

Figures 8–10 are the segmentation comparison results in terms of hippocampus, retinal blood vessel, sarcoma, and meningioma.

Given the segmented image, the IoU measure gives the similarity between the predicted region and the ground truth region for an object and is defined by following equation:

$$\text{IoU} = \frac{\text{TP}}{\text{FP} + \text{TP} + \text{FN}}, \quad (4)$$

where TP, FP, and FN denote the true positive, false positive, and false negative counts, respectively. The results are given in Table 6. We can see that the proposed segment method has the better result.

5. Conclusions

This paper proposes a medical image segmentation method based on multiscale convolutional neural network. This method can realize automatic segmentation of medical images and has high accuracy of segmentation. The CNN model in this paper not only reduces the amount of computation but also effectively captures multiscale information. In addition, the use of U-net fully mines the relevant information between slice sequences. Taking relevant medical image segmentation as an example, the experimental results on ADNI database show that the segmentation method in this paper is superior to other methods. The proposed method can perform segmentation tasks more easily and accurately. In the future, studying on deeply deep learning methods to segment images and applying them into different

types of images and different practical engineering are warranted.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] S. Yin and B. Jing, "Medical image annotation based on deep transfer learning," *Journal of Applied Science and Engineering*, vol. 22, no. 2, pp. 385–390, 2019.
- [2] L. Zhao, Z. Chen, Y. Yang, L. Zou, and Z. J. Wang, "ICFS clustering with multiple representatives for large data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 3, pp. 728–738, 2019.
- [3] L. Peng, Z. Chen, L. T. Yang, Q. Zhang, and J. Deen, "Deep convolutional computation model for feature learning on big data in Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 790–798, 2018.
- [4] Q. Zhang, C. Bai, L. T. Yang, Z. Chen, P. Li, and H. Yu, "A unified smart Chinese medicine framework for healthcare and medical services," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 129, pp. 1–8, 2019.
- [5] S. M. Mousavi, A. Naghsh, A. A. Manaf, and S. A. R. Abu-Bakar, "A robust medical image watermarking against salt and pepper noise for brain MRI images," *Multimedia Tools and Applications*, vol. 76, no. 7, pp. 10313–10342, 2017.
- [6] S. Kurtek, J. Su, C. Grimm, M. Vaughan, R. Sowell, and A. Srivastava, "Statistical analysis of manual segmentations of structures in medical images," *Computer Vision and Image Understanding*, vol. 117, no. 9, pp. 1036–1050, 2013.
- [7] L. Zhao, Z. Chen, L. T. Yang, M. Jamal Deen, and Z. Jane Wang, "Deep semantic mapping for heterogeneous multimedia transfer learning using co-occurrence data," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 15, no. 1, pp. 9–21, 2019.
- [8] X. Chen, J. Yao, Z. Ying, and U. Bagci, "3D automatic anatomy segmentation based on graph cut-oriented active appearance models," in *Proceedings of the IEEE International Conference on Image Processing*, Hong Kong, China, September 2010.
- [9] Q. Zhang, C. Bai, Z. Chen et al., "Deep learning models for diagnosing spleen and stomach diseases in smart Chinese medicine with cloud computing," *Concurrency and Computation: Practice and Experience*, p. e5252, 2019.
- [10] K. Y. Lim and R. Mandava, "A multi-phase semi-automatic approach for multisequence brain tumor image segmentation," *Expert Systems with Applications*, vol. 112, no. 1, pp. 288–300, 2018.
- [11] M. Nielsen, K. Petersen, M. Nielsen, and M. Lillholm, "A unifying framework for automatic and semi-automatic segmentation of vertebrae from radiographs using sample-driven active shape models," *Machine Vision & Applications*, vol. 24, no. 7, pp. 1421–1434, 2013.
- [12] L. Zhao, Z. Chen, Y. Yang, Z. Jane Wang, and V. C. M. Leung, "Incomplete multi-view clustering via deep semantic mapping," *Neurocomputing*, vol. 275, pp. 1053–1062, 2018.
- [13] M. G. Roberts, E. M. B. Pacheco, R. Mohankumar, T. F. Cootes, and J. E. Adams, "Detection of vertebral fractures in DXA VFA images using statistical models of appearance and a semi-automatic segmentation," *Osteoporosis International*, vol. 21, no. 12, pp. 2037–2046, 2010.
- [14] G. Wang, W. Li, M. A. Zuluaga et al., "Interactive medical image segmentation using deep learning with image-specific fine tuning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, pp. 1562–1573, 2018.
- [15] J. Gao, P. Li, and Z. Chen, "A canonical polyadic deep convolutional computation model for big data feature learning in Internet of Things," *Future Generation Computer Systems*, vol. 99, pp. 508–516, 2019.
- [16] 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," *Medical Physics*, vol. 43, no. 4, pp. 1882–1896, 2016.
- [17] 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 *Proceedings of the IEEE International Conference on Healthcare Informatics*, vol. 1, pp. 191–198, New York, NY, USA, June 2018.
- [18] W. Zhang, R. Li, H. Deng et al., "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," *NeuroImage*, vol. 108, pp. 214–224, 2015.
- [19] L. Zhao, Z. Chen, and Z. J. Wang, "Unsupervised multiview nonnegative correlated feature learning for data clustering," *IEEE Signal Processing Letters*, vol. 25, no. 1, pp. 60–64, 2018.
- [20] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Isgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1252–1261, 2016.
- [21] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, Las Vegas, NV, USA, June 2016.
- [22] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on International Conference on Machine Learning ICML'15*, vol. 37, pp. 448–456, Lille, France, July 2015.
- [23] P. Li, Z. Chen, L. T. Yang, J. Gao, Q. Zhang, and M. J. Deen, "An incremental deep convolutional computation model for feature learning on industrial big data," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1341–1349, 2019.
- [24] C. Liu, Z. Xiao, and N. Du, "Application of improved convolutional neural network in medical image segmentation," *Journal of Frontiers of Computer Science and Technology*, vol. 13, no. 9, pp. 1593–1603, 2019.
- [25] P. Pan, Y. Wang, Y. Luo, and J. Zhou, "Automatic segmentation of nasopharyngeal neoplas in MR image based on U-net model," *Journal of Computer Applications*, vol. 39, no. 4, pp. 1183–1188, 2019.
- [26] B. Gaonkar, D. Hovda, N. Martin, and L. Macyszyn, "Deep learning in the small sample size setting: cascaded feed forward neural networks for medical image segmentation," *Medical Imaging: Computer-Aided Diagnosis*, vol. 9785, 2016.
- [27] J. Gao, J. Li, and Y. Li, "Approximate event detection over multi-modal sensing data," *Journal of Combinatorial Optimization*, vol. 32, no. 4, pp. 1002–1016, 2016.

- [28] P. Li, Z. Chen, L. T. Yang, J. Gao, Q. Zhang, and M. J. Deen, "An improved stacked auto-encoder for network traffic flow classification," *IEEE Network*, vol. 32, no. 6, pp. 22–27, 2018.
- [29] R. C. Jack, M. A. Bernstein, N. C. Fox et al., "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 4, pp. 685–691, 2010.
- [30] D. Lu, K. Popuri, G. W. Ding, R. Balachandar, and M. F. Beg, "Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images," *Scientific Reports*, vol. 8, no. 1, p. 5697, 2018.
- [31] A. V. Oprobok, H. C. Achterberg, M. W. Vernooij, and M. De Bruijne, "Transfer learning for image segmentation by combining image weighting and kernel learning," *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 213–224, 2019.
- [32] F. Liu, "SUSAN: segment unannotated image structure using adversarial network," *Magnetic Resonance in Medicine*, vol. 81, no. 5, pp. 3330–3345, 2019.
- [33] Y. Wu, Y. Xia, Y. Song, Y. Zhang, and W. Cai, "Multiscale network followed network model for retinal vessel segmentation," in *Medical Image Computing and Computer Assisted Intervention-MICCAI*, pp. 119–126, Springer, Cham, Switzerland, 2018.