Uterine magnetic resonance image segmentation based on deep learning

Yu’ang Niu1,a, Yuanyang Zhang2,b, Liping Ying3,c, Hong Li4,d, Wenbo Chen5,e and Hanci Miao6,f, Nan Bao7,g*

1College of Medical and Biological Information Engineering, Northeastern University, Shenyang, LN 24, China
2College of Medical and Biological Information Engineering, Northeastern University, Shenyang, LN 24, China
3College of Medical and Biological Information Engineering, Northeastern University, Shenyang, LN 24, China
4College of Medical and Biological Information Engineering, Northeastern University, Shenyang, LN 24, China
5College of Medical and Biological Information Engineering, Northeastern University, Shenyang, LN 24, China
6College of Medical and Biological Information Engineering, Northeastern University, Shenyang, LN 24, China
7College of Medical and Biological Information Engineering, Northeastern University, Shenyang, LN 24, China

aemail: 614549600@qq.com, bemail: 648601617@qq.com, cemail: 524857137@qq.com, demail: lihong@bmie.neu.edu.cn, eemail: 1137256829@qq.com, femail: 1322216402@qq.com

*Corresponding author: semail: baonan@bmie.neu.edu.cn

Abstract. Endometrial diseases have always been a common disease among women worldwide. For the prevention of endometrial diseases and the treatment after illness, there is an urgent need to solve the problem of positioning the uterus, which is also one of the difficult problems that plague clinicians. Research in recent years has shown that neural networks based on deep learning are an important tool in the medical field. This article discusses the uterus magnetic resonance image segmentation method based on deep learning. First, we completed the preprocessing of the data based on the hessian matrix, and expanded the data, then inputted the data into the DenseUNet network. In this article, our method obtained 87.60% Dice Similarity Coefficient (DSC), 86.57% precision, 88.11% sensitivity and 99.75% specificity in 13 different test sets. Finally, our method was compared to UNet and UNet++ networks, and achieved better performance. Our method effectively solved the problems of uterus magnetic resonance image automatic segmentation. It can be used in the pre-operation planning of stem cell surgery to repair the endometrium.
1. Introduction

Endometrial diseases refers to various diseases that occur in the endometrial area, such as inflammation, injury, and cancer. It is one of the most common diseases in women. Endometrial diseases include endometritis, endometriosis, endometrial thickening, and endometrial cancer. Endometrial diseases can cause infertility in women of childbearing age. In severe cases, the uterus needs to be removed, and even life is endangered, causing heavy physical and mental pain to women. Therefore, the prevention and treatment of endometrial diseases is very important. Due to the unique biological characteristics of stem cells, it has great application prospects in the field of regenerative medicine and immunotherapy[1]. In recent years, regenerative medicine, which focuses on stem cell therapy, is expected to fundamentally repair the endometrium. However, in clinical practice, the lack of accurate positioning of the location of the endometrial injury has affected the clinical efficacy of stem cell transplantation and repair[2-3].

Magnetic Resonance Images (MRI) have good imaging capabilities for soft tissues, have very high resolution, and have a high SIGNAL-NOISE RATIO(SNR); different pulse sequences can be used to obtain multi-channel images with variable contrast, which can then be used for different target segmentation and classification of anatomical structures[4]. In modern clinical practice, MRI images are widely used in the differential diagnosis and clinical analysis of gynecological diseases[5]. In order to accurately locate the position of the endometrium, the segmentation of the uterine region based on the MRI image must be completed first. However, the traditional manual segmentation of the uterine region is a time-consuming and laborious task, and there are differences due to the subjective differences of the doctors. Therefore, it is very necessary to design a method that can automatically and accurately segment the uterine region.

With the continuous development of deep learning technology in recent years, deep learning has also begun to be applied in the medical field. Convolutional Neural Network(CNN) has shown excellent performance in many computer vision and machine learning problems, and it is also widely used in medical image segmentation[6].

Aiming at the problems of large storage and low computational efficiency in traditional segmentation methods based on CNN, Long Jonathan et al.[7] proposed the Fully Convolutional Networks (FCN) structure in 2015. FCN can classify images at the pixel level, solving the problem of semantic image segmentation.

In 2015, Olaf Ronneberger et al.[8] proposed the U-net network structure. The U-net network structure is similar to the FCN network structure. It is also divided into a down-sampling stage and an up-sampling stage. There are only convolutional layers and pooling layers in the network structure. In the connection layer, the shallow high-resolution layer in the network is used to solve the problem of pixel positioning, and the deep layer is used to solve the problem of pixel classification, so that image semantic level segmentation can be realized. U-net network performs well in the field of medical image segmentation. Many researchers use U-net network structure for medical image segmentation, and propose improvements on the basis of U-net network structure. Çiçek et al.[9] proposed a 3D U-net network structure, which realizes 3D image segmentation by inputting a continuous 2D slice sequence of 3D images. Milletari et al.[10] proposed a 3D structure V-net of U-net network structure. The V-net structure uses Dice coefficient loss function instead of traditional cross entropy loss function, and uses 3D convolution check image for convolution, and reduces channel dimension through $1 \times 1 \times 1$ convolution kernel. X Li et al.[11] proposed a new hybrid densely connected U-net (H-DenseUNet), which includes a 2D DenseUNet for effectively extracting on-chip features and a 3D DenseUNet counterpart segmentation for hierarchical aggregation. Among them, the on-chip and inter-chip features can be jointly optimized through the Hybrid Feature Fusion (HFF) layer.

We propose an automatic uterine cavity MRI segmentation method based on deep learning. In the experiment, the Hessian matrix is used to strengthen the uterine edge, then combines the enhanced image and the original image into the DenseUNet network. The segmentation accuracy of our method fully meets the standards required by doctors for surgery, helping doctors save a lot of time in finding the position of the uterus, and improve the efficiency and accuracy of surgery.
2. Methods

2.1. Uterine MR data preprocessing and enhancement

Hessian matrix has a wide range of applications in image processing, such as edge detection and feature point detection. The Hessian matrix is a square matrix composed of the second-order partial derivatives of a multivariate function. It describes the local curvature of the function and can be derived by using the Taylor expansion formula (1):

\[ \Delta f(x_0 + \Delta x) = f(x_0) + \Delta x \cdot f'(x_0) + \frac{1}{2!} \Delta x \cdot f''(x_0) \cdot \Delta x + o(\Delta x^2) \]

In a two-dimensional image, assuming that the function of image pixel value with respect to coordinates \((x, y)\) is \(f(x, y)\), then expand \(f(x + dx, y + dy)\) at \(f(x_0, y_0)\) to get as follows:

\[ f(x, y) \approx f(x_0, y_0) + \left[ \begin{array}{c} \Delta x \\ \Delta y \end{array} \right] \cdot \left[ \begin{array}{cc} f_x(x_0, y_0) & f_y(x_0, y_0) \\ f_y(x_0, y_0) & f_y(x_0, y_0) \end{array} \right] \cdot \left[ \begin{array}{c} \Delta x \\ \Delta y \end{array} \right] + \frac{1}{2!} \left[ \Delta x \Delta y \right] \cdot \left[ \begin{array}{cc} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{array} \right] \cdot \left[ \begin{array}{c} \Delta x \\ \Delta y \end{array} \right] \]

(2)

The second matrix in the third term on the right side of equation (2) is the Hessian matrix in the image; thus it is concluded that the Hessian matrix is the second derivative at one point of the image.

The second derivative of a point in the multidimensional space indicates how fast the gradient drops at that point. In a two-dimensional image, the Hessian matrix is a two-dimensional positive definite matrix with two eigenvalues and two corresponding eigenvectors. The two eigenvalues represent the anisotropy of the image change in the direction indicated by the two eigenvectors. According to the isotropy and anisotropy, the point structure and linear structure can be distinguished. In medical two-dimensional images, the second derivative is the degree of gray gradient change. The larger the second derivative, the less linear the gray change. The bright and dark string values indicate that the linear structure similar to the uterine edge is found from the classification according to the symbol of the feature value and the order of the feature vector and feature value. Table 1. indicates that the feature value may be positive (+) or negative (-), high (H) or low (L)[12]. They are respectively mapped to various structures, and the feature values mapped to the bright or dark linear structure are searched as indicators of potential detection of the uterine edge structure.

| Enhanced structure          | Hessian matrix parameters |
|----------------------------|---------------------------|
| Blank                      | \(\lambda_1\) | \(\lambda_2\) |
| Bright tubular structure   | L            | L           |
| Dark tubular structure     | L            | H-          |
| Bright blob-like structure | H-           | H-          |
| Dark blob-like structure   | H+           | H+          |

2.2. Uterine segmentation network design

With the deepening of network layers, the forward signal and the gradient signal in the network training process may gradually disappear after many layers. Previously, there were some very good ways to solve this problem. For example, in both Highway and ResNet structures[13], a kind of skip layer technology is proposed to make the signal flow between the input layer and the output layer at high speed. The core idea is to create a cross layer connection to connect the front and back layers of the network.

Densenet is a new connection mode based on this core concept[14-15]. In order to maximize the information flow between all layers in the network, all layers in the network are connected in pairs, so that each layer in the network accepts the characteristics of all layers in front of it as input.

In this project, the result of edge extraction is combined with the original image and sent to the
segmentation network to segment the uterus. The segmentation network adopts the design of 2D dense UNet, and its structure is shown in Figure 1. The depth of 2D DenseUNet extends to 167 layers, which is called 2D DenseUNet-167. It consists of 167 convolution layers, convergence layers, dense blocks, transition layers and up sampling layers. All the layers in the network are connected in pairs, so that each layer in the network accepts the characteristics of all the layers in front of it as input.

![Figure 1: Uterine segmentation network structure](image1)

The advantage of the network is to prevent the gradient near the input layer from becoming smaller and smaller with the increase of network depth, so it can reduce the gradient dissipation problem in the training process to a certain extent. Secondly, because a large number of features are reused, a large number of features can be generated by using a small number of convolution kernels, and the size of the final model is relatively small. The detailed structure of DenseUNet is shown in Figure 2.

![Figure 2: Denseblock schematic diagram](image2)

In order to change the size of the feature map, a transition layer is used, which is composed of batch normalization layer, $1 \times 1$ convolution layer and average pool layer. A compression factor is added to the transition layer to compress the number of feature graphs and prevent the expansion of feature graphs (set to 0.5 in the experiment). The upper sampling layer is implemented by bilinear interpolation, followed by the summation of lower layer features (i.e. u-net connection) and $3 \times 3$ convolution layers. Before each build-up layer, a batch normalized and corrected linear unit (ReLU) is used[16].

Before training, all convolutional layers are initialized by Xavier, and Stochastic Gradient Descent
(SGD) is used to update the weights in the network[17]. For the optimization of the network, we use the Adam optimizer and set the initial learning rate to 0.01. During the training process, the learning rate decays according to the formula (3).

\[ \text{decayed}_{\text{lr}} = \text{initial}_{\text{lr}} \times \frac{\text{global}_{\text{steps}}}{\text{decay}_{\text{steps}}} \] (3)

Among them, \text{initial}_{\text{lr}} represents the initial learning rate, \text{decay}_{\text{rate}} represents the decay rate, \text{global}_{\text{steps}} represents the current training round, \text{decay}_{\text{steps}} defines the decay period, the \text{batchsize} is set to 4, and the hyperparameters are set to \text{epsilon}=1e-3 and \text{momentum}=0.99.

The loss function is calculated according to the following formula (4).

\[ L(y, y^\prime) = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C} w_i^c y^c_i \log \hat{y}_i^c \] (4)

\( \hat{y}_i^c \) denotes the probability of voxel \( i \) belongs to class \( c \), \( w_i^c \) denotes the weight and \( y_i^c \) indicates the ground truth label for voxel \( i \).

3. Results

3.1. Data preparation

This project obtained a total of 77 MRI data of patients undergoing pelvic MRI imaging examinations at Shengjing Hospital affiliated to China Medical University from December 31, 2015 to May 27, 2019. The doctor completes the labeling of the original data. Due to the different specifications and sizes of the original clinical data, we first uniformly crop or fill the image size to 512*512, and perform unified spatial normalization, histogram matching, and gray-scale normalization on all MR images. The shape of the uterine edge is a linear structure, so we use the Hessian matrix to preprocess the data to enhance the linear structure, as shown in Figure 3(b). Combining the enhanced data with the original data for deep learning training will improve the ability to recognize the uterine boundary.

![Figure 3 Raw data and boundary enhancement results](image)

Next, in order to solve the problem of too little training data for deep learning networks, use data expansion (vertical and horizontal flip, 20°rotation, x-axis displacement 10%, y-axis displacement 10%, scale down to 85%, zoom in to 115 %, affine transformation 17°) enlarges the original data volume by 12 times, and the effect of the corresponding preprocessing is shown in Figure 4.
3.2. Evaluation of segmentation results

We conducted training and testing on MR images of 50 and 13 patients respectively. The number of images in the test set is 260. We performed a five-fold cross-validation on the original hospital data, four-fifths of the data were used for training, and one-fifth of the data was used for verification. The method we propose is based on the Python3.6.2 and Pytorch1.6.0. The computer used in the experiment consists of Intel Xeon E5-2640 2.40GHz CPU, 16GB RAM and NVIDIA1080ti GPU. The operating system is windows10.

Figure 5 shows the effect of our method to segment the uterus. The part marked in green is the uterine area. After the doctor's confirmation, our results can fully meet the standard. Next is our quantitative analysis.

Table 2 shows the results of our method under four different evaluation criteria on 13 test sets. Our method achieved an average of 87.60% DSC, 86.57% precision, 88.11% sensitivity and 99.75% specificity in 13 test sets.

| Test Datasets | DSC    | Precision | Se     | Sp     |
|---------------|--------|-----------|--------|--------|
| 1             | 88.79% | 87.63%    | 89.58% | 99.73% |
| 2             | 89.27% | 85.09%    | 94.71% | 99.79% |
| 3             | 83.85% | 90.39%    | 77.34% | 99.70% |
4. Conclusions

Our project addresses the problem of the difficulty in accurately positioning the uterus in the endometrial repair surgery. First, in the original data processing stage, the edges extracted by the Hessian matrix are combined with the original data to achieve data enhancement; at the same time, the amount of data is increased through the basic transformation of the image; then the processed data is sent to the DenseUNet network. The experimental results show that our method has quite good results in many different data sets, and compared with the recognized advanced network in the field of medical image processing, our method has the best performance in various evaluation criteria. Various results prove that our method can meet the requirements of clinicians for preoperative planning of endometrial repair surgery. It has very important application prospects and values to promote the development and application of stem cell precision repair surgery.

Due to conditions, this project still has some shortcomings, such as relatively few experimental data. In the future work, we can obtain more data through strengthened cooperation with hospitals. At the same time, we will try to design better networks, such as cascading multiple networks, and try to use better data, such as multimodal uterine images.

Acknowledgements

The authors would like to thank the editor and reviewers for improving the quality of the paper. This work was partially supported by National Natural Science Foundation of China (NSFC) [Grant No.61873257, 61701103], National Natural Science Foundation of Liaoning Province [Grant No.2019-ZD-0005].

References

[1] Nagori C B, Panchal S Y, Patel H. (2011) Endometrial regeneration using autologous adult stem cells followed by conception by in vitro fertilization in a patient of severe Asherman's syndrome[J]. Journal of human reproductive sciences, 4(1): 43.

[2] Bardou, B., Nageotte, F., Zanne, P., & De Mathelin, M. (2012) Improvements in the control of a flexible endoscopic system. IEEE International Conference on Robotics & Automation. IEEE.
[3] Sylvestre C, Child T J, Tulandi T. (2003) A prospective study to evaluate the efficacy of two-and three-dimensional sonohysterography in women with intrauterine lesions[J]. Fertility and Sterility, 79(5): 1222-1225.

[4] Vandenberghhe S, Marsden P K. (2015) PET-MRI: a review of challenges and solutions in the development of integrated multimodality imaging[J]. Physics in Medicine & Biology, 60(4):115-54.

[5] Y Kurata, M Nishio, A Kido, K Fujimoto, K Togashi. (2019) Automatic segmentation of the uterus on MRI using a convolutional neural network[J]. Computers in Biology and Medicine, 114:103438-.

[6] Krizhevsky, A., Sutskever, I., & Hinton, G. (2012) ImageNet Classification with Deep Convolutional Neural Networks. NIPS (Vol.25). Curran Associates Inc.

[7] Long J, Shelhamer E, Darrell T. (2015) Fully Convolutional Networks for Semantic Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(4):640-651.

[8] Ronneberger O, Fischer P, Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation[J].

[9] Targ, S., Almeida, D., & Lyman, K. (2016) Resnet in Resnet: Generalizing Residual Architectures. ICLRW.

[10] Huang, G., Sun, Y., Liu, Z., Sedra, D., & Weinberger, K. Q. (2016) Densely connected convolutional networks[J].

[11] Nair, V., & Hinton, G. E. (2010) Rectified Linear Units Improve Restricted Boltzmann Machines. Proceedings of the 27th International Conference on Machine Learning (ICML-10), June 21-24, 2010, Haifa, Israel.

[12] Gardner, W. A. (1984) Learning characteristics of stochastic-gradient-descent algorithms: a general study, analysis, and critique. Signal Processing, 6(2), 113-133.