ISSN: 2664-9640

Overview of Colorectal Cancer Detection Based on U-Net

Mingao Liu

University of Science and Technology Liaoning, Anshan 114000, China.

Abstract

Using computer vision technology for pathological diagnosis has become a hot spot in the medical field. The progress of this kind of research helps to alleviate the pressure of doctors and effectively save medical resources. But there is still much room for improvement in current technologies. Accurate segmentation of medical images has a crucial impact on the evaluation and treatment of the disease. Since the difference between cancer cells and surrounding tissues in medical images is not obvious, it increases the difficulty of segmentation. In this paper, the U-Net network structure is used to solve the medical image segmentation problem with the detection of rectal cancer as the goal.

Keywords

U-Net ; semantic model ; rectal cancer ; network model.

1. Introduction

Cancer is now known as one of the difficult to cure diseases, making people talk about cancer discoloration. Among the many types of cancer, rectal cancer has become one of the more common cancers. Every year, people worldwide lose their lives because of rectal cancer, and its incidence is also high, accounting for one tenth of all cancers. There are also countless people found with rectal cancer in China. Among them, about 388,000 new cases of rectal cancer occur every year, the incidence ranks the third among all malignant tumors, and the mortality rate ranks the fifth, reaching 187,000[1] . For doctors who need to observe many pathological slices every day, it is impossible to keep a high degree of attention at all times under high pressure. In order to improve the accuracy of the diagnosis of pathological slices, it is necessary to use artificial intelligence to judge pathological images. While reducing the workload of doctors, it improves the efficiency of diagnosis of diseases and greatly saves medical resources. However, how to improve the accuracy of segmentation is still a major challenge in the field of medical images.

The emergence of deep learning has greatly changed the research on image vision. As a classical application structure in deep learning, convolutional neural network (CNN) is composed of convolution layer, pooling layer and fully connected layer. This structure can automatically extract features from large amounts of data, and has good performance in image segmentation, target recognition and other tasks. Image semantic segmentation is to assign a semantic category label to each pixel in the image, which is the classification of image pixel level. Long et al. (2015) [2] proposed a fully convolutional neural network (FCN) based on semantic segmentation, which used convolution layer to replace the full connection layer on the basis of CNN.Ronneberger et al. (2015) [3] proposed the U-Net network with end-to-end encoder-decoder structure, which is improved based on FCN. The emergence of U-Net network overcomes the shortcomings of FCN pixel space location information can not be fully retained and the loss of local and global features caused by context information. Zhou et al. (2018) [4] proposed the U-Net + + network, modified the layers of the U-Net network, redesigned the jump connection, and used deep supervision to use different networks for different segmentation tasks to achieve the best results.

Based on the pathological images of rectal cancer, this paper uses computer vision and other technologies to train them to improve the accuracy, and uses U-Net[5] to complete the experiment. Through the training of the model to achieve the best effect, assist doctors in the diagnosis of pathological sections. Therefore, the detection of nuclei in pathological images has become a top priority and also one of the difficulties in this paper.

2. U-Net

Ronneberger et al. (2015)[3] proposed the U-Net network. U-Net has sufficient influence on the semantic segmentation of medical images. In the few years since its publication, the reference amount is surprisingly high. The U-Net structure uses the U-shaped network structure to obtain the context information and location information. The coding structure of U-Net is down-sampled four times, a total of 16 times. On the contrary, the decoding structure is also up-sampled four times, and the image is restored. While upsampling, skip connection is used at the same stage to ensure better fusion of restored feature maps. Four up-samplings made the information such as edge recovery more detailed after segmentation.

2.1. Encoder Improvement

U-Net； The U-Net encoder extracts the feature of the image, and the decoder restores the size of the extracted feature map. There are many improvements to the encoder of the U-Net model, such as Y-Net, Ψ-Net and multipath dense U-Net ;

Lan et al. (2020) [6] proposed the Y-Net network. The difference between this model and the original model is that the network structure used is the‘Y’structure, which is composed of two encoders and a decoder. The two encoders can make the obtained feature information more accurate and detailed. The input images are typed into two encoders respectively. The second encoder delays the addition of the input information when input. Encoder I processes the input images and connects the output of the two encoders to generate the final segmentation results.

Kuang et al. (2020) [7] proposed Ψ-Net, its network presents the Greek alphabet ' Ψ ' type, its network structure has three encoders and a decoder. The encoder processes the slices that need to be processed and the adjacent slices, using a self-attention block. Context attention blocks are used at the decoding layer. This design helps to extract global features and repair images with low contrast to achieve the accuracy of segmentation. Mainly used in irregular shape and location of the brain hemorrhage image segmentation.

Dolz et al. (2019) [8] proposed multi-path dense U-Net. In order to extract the characteristics of different brain imaging, the input of the coding path includes diffusion-weighted imaging (DWI), cerebral blood volume (CBV), CT perfusion imaging (CTP), and mean transit time (MMT). Full use of the characteristics of various modes in the coding end can prevent the disappearance of the gradient, due to the reference form and small reduction of overfitting.

2.2. Model Combination

The integration of U-Net network with other models will produce better results, and the combination of different modules in different areas can greatly improve the results in this direction. There are many modules that are combined with it, such as deep reinforcement learning, graph cut method, local difference method, CNN and so on.

Man et al. (2019) [9] proposed a deformable U-Net network driven by deep reinforcement learning, which mainly solves the problem of complex background and difficult positioning. Firstly, the required part of the image is clipped by reinforcement learning, and then the processed image is segmented by deformable convolution. Liu et al. (2019) [10] proposed a liver segmentation method combining U-Net and graph cut method. The model is mainly for the diagnosis of liver and nearby organs. Firstly, the positive liver of the image is segmented to form a probability distribution graph, and then the segmentation is completed by minimizing the

graph cut energy function using the context information. Zhang et al. (2020d) [11] applied the combination of U-Net and local difference method to the identification of cerebral hemorrhage images. Firstly, the information in the image is filtered by the threshold segmentation algorithm, and the feature is extracted to locate the bleeding point, and then the image is segmented by U-Net to achieve the purpose.Wu et al. (2020) [12] proposed a composite model combining CNN and U-Net to accurately segment the left ventricle of MRI images in view of the difficulty in left ventricle segmentation in MRI images.

2.3. 3D U-Net Network

3D images show more information than 2D images. The more information displayed in medical diagnosis, the more accurate the diagnosis of patients will be. If 2D U-Net model is used to process 3D images, the calculation will be stupid and expensive.

Hu et al. (2018) [13] proposed a 2.5D segmentation network based on U-Net. Three 2D U-Net networks are used to reintegrate the predicted probability maps of three different directions, and the final results are obtained. However, the efficiency of this 3D image segmentation method is too low.。 Man et al. (2019) [9] Segmentation of coronal, sagittal and transverse sections of pancreatic patients using 2D U-Net networks composed of deformable convolutions. The final segmentation results of sliced images in three directions stack together.Çiçek et al. (2016) proposed 3D U-Net, extending the 2D U-Net network to 3D spaces, including analytical paths for abstract features and synthetic paths for generating segmentation, and establishing shortcuts between the two paths. Milletari et al. (2016) [15] proposed V-Net volume convolution neural network. Dou et al. (2017) proposed an improved 3D full convolution network, using 3D monitoring mechanism to accelerate the contraction of deep network in the case of insufficient training set.

2.4. Structure Modification

The improvement of U-Net structure includes the improvement of data enhancement, convolution operation, down-sampling operation, up-sampling operation, model optimization strategy and jump connection. Data enhancement has a variety of operations, such as changing the size of the input image, adding noise, etc. The ultimate goal of data operation is to increase data samples, which can better train the model.。 For the problem of limited medical samples of U-Net network, elastic deformation is used for data enhancement, including elastic transform, geometric transformations, generative adversarial network (GAN), WGAN (Wasserstein GAN) and real-time intensifier.Convolution operations extract features from the original image. The U-Net network needs to use activation function after convolution operation to normalize the results to solve the problem of gradient disappearance in the model running direction. Therefore, the improvement of convolution operation includes convolution block and convolution filling method. The up-sampling operation is the operation of feature extraction to restore the feature map to the original size. There are mainly five commonly used methods : transpose convolution, nearest neighbor interpolation, bilinear interpolation, trilinear interpolation and sub-pixel convolution.。 After convolution, the U-Net network performs the maximum pooling operation to further select the features. The improvement of pooling layer mainly includes seven methods : maximum pooling, average pooling, random pooling, span convolution, dilation convolution, spatial pyramid pooling and Inception module. The function of jump connection in U-Net network is to prune the convolution feature map and add it to the decoding to realize the positioning of pixels. Jump connection methods include attention mechanism block, feature reuse and attention mechanism block (FRAM), deconvolution + activation function, annotation information obtained by Siam network and new jump connection methods.

3. Conclusion

Firstly, this paper expounds the basic structure, working principle and model improvement method of U-Net network. Then summarize the structure optimization methods of U-Net network ; finally, the mechanisms of network improvement are reviewed, including residual thinking, intensive thinking, attention mechanism and the combination of multiple mechanisms, which provides reference for the diagnosis and treatment of clinicians. In recent years, there have been studies on the U-Net network in the field of medical images, such as the diagnosis of liver and nearby organs, the detection and segmentation of left ventricular images, and the detection and recognition of cerebral hemorrhage in CT images. We have analyzed the research that is similar to the detection task of rectal cancer images. Finally, we can obtain the effectiveness of the U-Net network model in the detection of rectal cancer.

Acknowledgements

Liaoning University of Science and Technology College Students Innovation and Entrepreneurship Training Project Fund(CT image rectal cancer detection system based on Unet + +).

References

[1] Su Wenyu, and Fang Jingyuan. ' New progress in colorectal cancer research in 2020. ' Chinese Medical Information Report 36.3 (2021) : 2.
[2] Shelhamer, E. , J. Long , and T. Darrell . "Fully Convolutional Networks for Semantic Segmentation." IEEE Transactions on Pattern Analysis and Machine Intelligence (2016):1-1.
[3] Ronneberger, O. , P. Fischer , and T. Brox . "U-Net: Convolutional Networks for Biomedical Image Segmentation." Springer, Cham (2015).
[4] Zhou, Z. , et al. "UNet++: A Nested U-Net Architecture for Medical Image Segmentation." (2018).
[5] Han, Y. , and J. C. Ye . "Framing U-Net via Deep Convolutional Framelets: Application to Sparse-View CT." IEEE Transactions on Medical Imaging (2018):1418.
[6] Lan, H. , et al. "Y-Net: Hybrid Deep Learning Image Reconstruction for Photoacoustic Tomography in vivo." Photoacoustics 20(2020):100197.
[7] Kuang, Z. , et al. "Ψ-Net: Focusing on the border areas of intracerebral hemorrhage on CT images." Computer Methods and Programs in Biomedicine 194(2020):105546.
[8] Dolz, Jose , I. B. Ayed , and C. Desrosiers . Dense Multi-path U-Net for Ischemic Stroke Lesion Segmentation in Multiple Image Modalities: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part I. 2019.
[9] Man, Yunze , et al. "Deep Q Learning Driven CT Pancreas Segmentation with Geometry-Aware U-Net." (2018).
[10] Liu, Y. , et al. "CR-U-Net: Cascaded U-Net with Residual Mapping for Liver Segmentation in CT Images *." 2019 IEEE Visual Communications and Image Processing (VCIP) IEEE, 2019.
[11] Zhang Tianqi et al. ' Intracranial hemorrhage recognition algorithm based on U-Net. ' Journal of Beijing University of Posts and Telecommunications 3 (2020) : 7.
[12] Wu, B. , Y Fang, and X. Lai . "Left ventricle automatic segmentation in cardiac MRI using a combined CNN and U-net approach." Computerized Medical Imaging and Graphics 82.
[13] Ke, H. , et al. "A 2.5D Cancer Segmentation for MRI Images Based on U-Net." 2018 5th International Conference on Information Science and Control Engineering (ICISCE) IEEE, 2018.
[14]iek, zgün, et al. "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation." Springer, Cham (2016).
[15] F Milletari, N. Navab , and S. A. Ahmadi . "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation." IEEE (2016).
[16] Dou, Q. , et al. "3D deeply supervised network for automated segmentation of volumetric medical images." Medical Image Analysis (2017).