ISSN: 1813-4890

CT Image Segmentation based on UNET ++ and UNET Network

Mingao Liu, Ping Zou

University of Science and Technology Liaoning, Anshan, 114051, China

Abstract

Pancreatic cancer, as one of the common malignant tumors, causes a large number of deaths every year. For this reason, electronic sectional examination (CT) is used as a screening tool for pancreatic cancer diagnosis. In the pancreas CT image, medical experts can distinguish whether the tumor inside the pancreas is benign or malignant, which is of great significance for guiding clinical treatment. To solve this kind of problem, use Unet++ to solve the problem of image segmentation.

Keywords

Unet++; Unet; Image Segmentation.

1. Introduction

With the improvement of social and economic levels, people's diet structure has undergone great changes. Pancreatic cancer, as one of the most common malignant tumors in China, has been widely concerned in China. Among the 1.36 million cases of pancreatic cancer diagnosed worldwide in 2012, nearly 300,000 new cases occurred in China, accounting for one-fifth of the new cases worldwide. As China has the largest number of new cases of pancreatic cancer in the world, active cancer prevention and treatment are of great significance to reduce the incidence and mortality of colorectal cancer in China. With the improvement of people's living standards and the change in lifestyle, the incidence of colon cancer is increasing year by year, but the number of pancreatic cancer patients in China still accounts for the majority of pancreatic cancer[1]. Especially in the aging society, the incidence and correlation of pancreatic cancer, one of the most common cancers, are becoming more and more important[2]. Major risk factors include advanced age, family history, male sex, and lifestyle factors. Proper cancer screening can reduce the incidence and mortality of pancreatic cancer.

With the development of the social economy, deep learning algorithm has been recognized by the public, and medical image detection and image segmentation algorithm based on deep learning has been gradually used by the public. The deep learning framework can automatically identify whether there are tumor regions in CT images and accurately segment the tumor regions. Therefore, it is of great practical significance to make full use of modern information technology to build an algorithm model of pancreatic cancer image tumor recognition and segmentation based on a convolutional neural network, which greatly reduces the workload of medical experts and saves a lot of time for subsequent diagnosis and treatment.

2. Technical

This chapter introduces relevant theories and technologies this paper. First, it describes common medical image recognition and segmentation methods and then focuses on deep learning technologies, including an overview of deep learning and the basic structure of the convolutional neural network. Finally, it summarizes this section.

2.1. Medical Image Segmentation Method based on Deep Learning

(1) FCN network

ISSN: 1813-4890

With the rapid development of deep learning, medical image segmentation based on a deep convolutional neural network (CNN) has become a research hotspot. The traditional CNN network extracts features through the convolutional layer and outputs results through the subsequent full connection layer, but the results are all a category probability value of the input image. This network structure is very suitable for image classification tasks. But the image segmentation task is different from the classification task, the key point of this task is how to accurately classify each pixel in the image to achieve the purpose of segmentation. To achieve image segmentation technology, Long[3]Et al. proposed a full convolutional neural network (FCN).

The network is based on the traditional convolutional neural network to make the following improvements:

1. Replace all the internal full connection layers with the convolution layer.

2. Add a deconvolution layer. The input image is firstly extracted through the convolutional layer and pooling layer of FCN, and then the obtained feature image is up-sampled through the deconvolution layer to restore it to the same size as the input image. Finally, each pixel is classified to achieve the segmentation purpose.

FCN network will extend the image classification technology to image segmentation technology, the use of end-to-end operation to achieve the classification of image pixels, improve the segmentation effect, for the future image segmentation technology laid a good foundation.

(2) Unet network

For medical images, the FCN network does not take into account the global context information, resulting in unsatisfactory segmentation details. To solve the above problems, Ronneberger [4] Et al. proposed the Unet network model.

The model is based on the encoder-decoder method, the encoder is mainly used to extract the context information of the image through several convolution and down-sampling operations, and the decoder is used to accurately locate the target through several convolution and up-sampling operations. The biggest characteristic of the network is to join the skipped path connection, which closely connects the nodes at the same level in the encoder and decoder, and realizes the enhancement of data information. Because Unet makes full use of the context information of images, it is widely used in medical image processing tasks. According to reading a large number of papers related to medical image segmentation, it is found that most of the networks used are improved based on Unet.

2.2. Convolutional Neural Network

As one of the popular neural networks in deep learning technology, the convolutional neural network has been widely used in the field of medical image processing, such as image recognition, image segmentation, and target detection, and has achieved good results. The neural network consists of several convolutional layers, pooling layers, and full connection layers. Next, this paper will focus on the basic structure of the convolutional neural network.

(1) Convolution layer

The convolutional layer is the most important part of a convolutional neural network. Its function is to realize feature extraction of images through convolution operation between the inner convolution kernel and the input image. Each parameter of each convolution kernel in the convolution layer corresponds to the weight value and bias value in the network. Since a single convolution kernel extracts the local features of the input image, the size of the convolution kernel determines the size of the local feature graph. And the convolution operation.

The most important feature of the convolutional neural network is to realize internal local connection and parameter sharing through the convolutional layer. Local connection refers to the connection between the output neurons in the network and some input neurons through

ISSN: 1813-4890

the convolution operation, which reduces the number of parameters in the network and speeds up the training time of the network model. Parameter sharing, also called weight sharing, means that the same convolution kernel can extract the features of different regions, and this extraction method is independent of the location of the region. This method greatly reduces the amount of network computation and improves the efficiency of network learning.

(2) Pooling layer

After the input image passes through the convolutional layer, a large amount of feature information will be extracted. If all this information is put into the output layer for classification, the following problems will occur:

1. Processing a large amount of data will inevitably increase the time of network operation, and has a great demand on hardware conditions.

2. It's easy to overfit. The extracted feature information contains irrelevant information. If the network learns the redundant feature information, it will inevitably lead to the over-fitting phenomenon of the network.

To solve these problems, convolutional neural networks add a pooling layer after the convolutional layer. The pooling layer, also known as the down-sampling layer, is used to reduce the data dimension of image features extracted from the convolution layer, filter redundant feature information, and then extract important feature information in the image. In this way, the computation amount of the network is greatly reduced, the computational efficiency of the network is accelerated, and the problem of over-fitting the network is avoided.

The pooling operation can be divided into average pooling and maximum pooling. Through reading the latest literature, we know that maximum pooling is the most commonly used method at present, which is generally better than average pooling in most image classification and segmentation tasks. Pooling operation refers to dividing the input feature map into several equal regions. Maximum pooling retains the largest element in the region as the new feature map, while average pooling averages the elements in the region to form a new feature map.

(3) The connection layer

The full connection layer mainly appears in the image classification network. In general, it is located after the convolution layer and the pooling layer. Its function is to connect the extracted image features and finally realize the classification task. In the full connection layer, each output neuron will connect with all input neurons to achieve the effect of feature combination. This kind of connection is very similar to the ordinary neural network.

3. Unet++ Network

According to the structure of Unet network, as the Unit network starts to conduct upsampling operations after 4 downsampling times, the Unet network extracts more deep features of images. For an image, its internal shallow features and deep features are equally important. Therefore, to make up for the shortcomings of the Unet network, Zhou 2018 [5] Et al. proposed the Unet++ network model, whose network structure is shown in Figure 4.1, which is improved based on Unet network model. First, the Unet++ network carries out an upsampling after each downsampling of the Unet network. Then, inspired by the Densenet network, we redesigned a series of nested, dense skipped path connections. Its purpose is to reduce the semantic gap between the feature map information.

The Unet++ network model is also based on the encoder-decoder structure. The black part represents the original U-NET network, and the green part represents the dense convolution block in the skipped path in the middle of the Unet network. The connectivity of the encoder, as well as the decoder part of the Unet++ network, has been changed due to the redesign of the skipped path connection. For Unet networks, the encoder part transmits feature information to

ISSN: 1813-4890

the decoder part by skipping between path connections. In Unet++, the features extracted by the encoder will be transmitted to the dense convolution block in the middle for feature extraction again. For example, node $X^{1,0}$And node $X^{2,2}$The skipped path between two convolution layers consists of two dense convolution blocks. In front of each convolution layer is the connection layer, whose function is to fuse the output of the previous convolution of the current dense convolution block with the output of the upper sampling of the lower dense convolution block. In essence, the function of a dense convolution block is to reduce the semantic gap between encoder and decoder feature mapping. The performance of the model is improved when the feature maps of the encoder and the corresponding decoder are semantically similar.

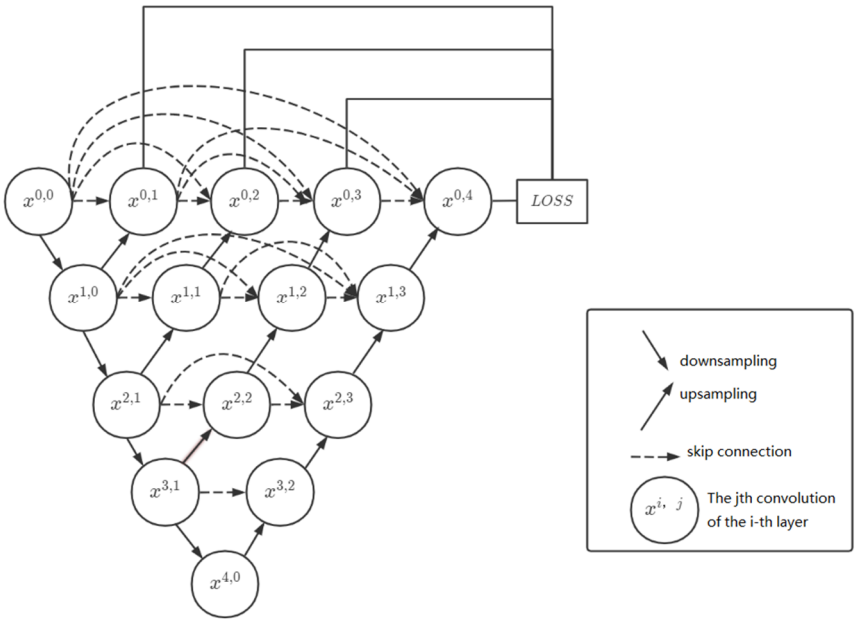

Figure 1. Unet++ network structure

The Unet++ network is equivalent to connecting Unet networks at layer 1 to layer 4 in the form of long and short connections. This connection can extract all the depth features so that the network can learn the important features in the image itself.

4. Summarizes

This paper introduces several algorithms for the recognition and segmentation of pancreatic cancer CT images. In traditional cancer screening, medical experts use manual screening to screen pancreatic CT images for tumors. However, after medical experts read a large number of CT images, there will be subjectivity and visual fatigue. This paper uses a deep learning-based segmentation algorithm for pancreatic tumors to assist doctors in diagnosis. By investigating different deep learning algorithms and comparing algorithms, this paper finally uses the Unet++ network to solve the CT image of pancreatic cancer. Pancreatic tumor is one of the most deadly malignant tumors in the world, and accurate segmentation of pancreatic tumors from CT images is particularly important.

Acknowledgments

Liaoning University of science and technology innovation and entrepreneurship training project fund (CT image rectal cancer detection system based on unet++).

ISSN: 1813-4890

References

[1] Peng Jun-jie, ZHU Ji, LIU Fang-qi, Chen Zhi-yu, Tong Tong, HUANG Dan, LI Wen-hua, Yang Li-feng. Chinese journal of cancer,2017,27(01):41-80.

[2] Bogner A , Kirchberg J , Weitz J , et al. State of the Art Rectal Cancer Surgery[J].Visceral Medicine, 2019, 35(4):1-7.

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

[4] Ronneberger O, Fischer P , Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation [J]. Springer International Publishing, 2015.

[5] Zhou Z, Siddiquee M , Tajbakhsh N , et al. UNet++: A Nested U-Net Architecture for Medical Image Segmentation [J]. 2018.