Deep Learning for Image Segmentation

Xueting Zhou, Yan Chen and Shoushan Liu

October 4, 2022

Deep learning for image segmentation

Ultrasound image segmentation of thyroid nodules based on U_Net

Xueting Zhou

Shandong University of Science and Technology,Qingdao,China,202182130042@sdust.edu.cn

Yan Chen

Shandong University of Science and Technology,Qingdao,China,1363chen@163.com

Shoushan Liu

Shandong University of Science and Technology,Qingdao,China,lsscy@hotmail.com

The purpose of this article is to explore the application value of deep learning algorithm in ultrasound images of thyroid nodules. A dataset of 7288 ultrasound images of thyroid nodules collected from the MICCAI 2020 challenge is established, the U_Net method is used, and adjustments are made on the basis of this method. Through continuous training, the optimal model is found and the computation is autonomous. accurate segmentation of thyroid nodules. The segmentation accuracy obtained by this network reaches 0.955, which has good segmentation performance.

Additional Keywords and Phrases: Deep learning, U_Net, Thyroid nodules

1 INTRODUCTION

1.1 Research background and significance

The existence of thyroid nodules is generally not easy to detect, but malignant nodules can cause the thyroid to secrete thyroid hormone abnormally, which in turn affects blood pressure, heart rate, body temperature, etc., and can lead to death in severe cases.

Generally,the position and aspect ratio of nodules need to be manually calibrated by doctors, which is time-consuming, labor-intensive and subjective. With the development of science and technology, artificial intelligence technology has been applied in the medical field, and deep learning methods have achieved great results in medical image processing. Computer-aided diagnosis can reduce the burden of the doctor, saves manpower, and also improves the accuracy and efficiency of the diagnosis.

The computer-aided diagnosis of thyroid nodules has gone through about three stages. The first stage is the expert system method, but it is not intelligent because it contains a large number of if-then-else statements and requires manual intervention. The second stage is the artificial feature method and machine learning method, that is, the features of thyroid nodules are manually extracted and applied to the image, but the shapes of thyroid nodules are different and easily mutated.The third stage is deep learning methods. With the development and introduction of deep learning, breakthroughs have been made in the field of computer vision, and convolutional neural networks have become an important tool for image processing, as they can make full use of the deep feature information of images to complete the semantic segmentation task of images. A series of deep learning based image segmentation methods have been proposed commonly used segmentation methods such as FCN, U_Net and PSPNet[1].

1.2 Research status at home and abroad

In the procedure of Computer-aided diagnosis of thyroid nodules based on ultrasound images, segmentation of the thyroid nodules always determines the results of thyroid nodules classification.Segmentation of thyroid ultrasound images into nodules and other regions based on the semantic characteristics of thyroid ultrasound images provides a prerequisite for subsequent benign and malignant classification.Many researches over the years have promoted the development of image segmentation technology.

Bosch [2] proposed the random number forest method for classifiers in 2007, which immediately became a very popular machine learning method. Jonathan Long [3] researched and invented the Fully Convolutional Neural Network (FCN) in 2015. The network mainly includes three parts, full convolution, deconvolution, and layer structure. FCN replaces the last fully connected layer of CNN with a convolutional layer. Ronneberger [4] proposed a U_Net model based on FCN, which has a great promise in image segmentation. JinLian Ma [5] applied convolutional neural network in thyroid ultrasound image segmentation for the first time. He trained 6,000 ultrasound images of thyroid nodules and validated 15,000 images. He found that the convolutional neural network was effective in segmenting thyroid nodules, and then improved it to obtain an effective ultrasound segmentation model for thyroid nodules, achieving a relatively high accuracy rate.

1.3 The research content of this paper

In this paper,an experiment using the dataset of the MICCAI 2020 challenge is carried out. Since the original dataset is not published , only the enhanced dataset is downloaded. The enhanced dataset has a total of 7288 ultrasound images of thyroid nodules. The image is based on the original image and turns one image into two by rotating, translating, and adding noise. An image of segmented thyroid nodules corresponding to the original image is also provided for training in deep learning. The 3644*2 dataset in this experiment is divided into training set, validation set and test set according to the ratio of 7:2:1.

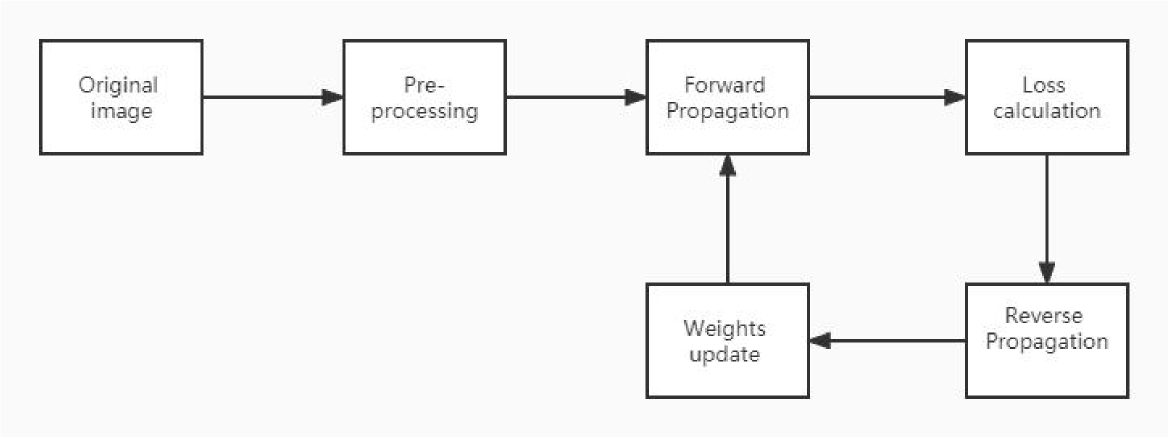

Figure 1: Flow chart of thyroid nodule segmentation

As shown in Figure 1, this experiment is roughly divided into the following steps: the first step is to pre-process the image,firstly, the image is desensitized to delete the large borders and some useless information in the thyroid ultrasound image, and then the histogram equalization method is used to equalize and smooth the image to avoid the image being too bright or too dark; the second step is the segmentation of the nodules, forward propagation, and then calculate the loss for back propagation, update the parameters, and save the parameters.In addition to data pre-processing, the latter process is in a loop until the training is completed by a certain number of training; the third step is to use the model trained by the deep convolutional neural network to test the accuracy of the model.

2 EXPERIMENTAL PRINCIPLES

As shown in Figure 2,the U_Net network structure consists of three parts, namely up-sampling, down-sampling and skip connections. It can also be divided into two parts: the left part is the encoding part, where the image size will be compressed to obtain some relatively simple features of the nodules. On the right is the decoding part, where some deeper nodule features are extracted. After each convolution, the image size will become smaller, because the convolution is filled in a valid way to ensure that the result of the convolution is obtained without missing context. In the middle, the feature map obtained by encoding and the feature map obtained by decoding are combined together to refine the image. Finally, the segmentation prediction is performed according to the obtained feature map.

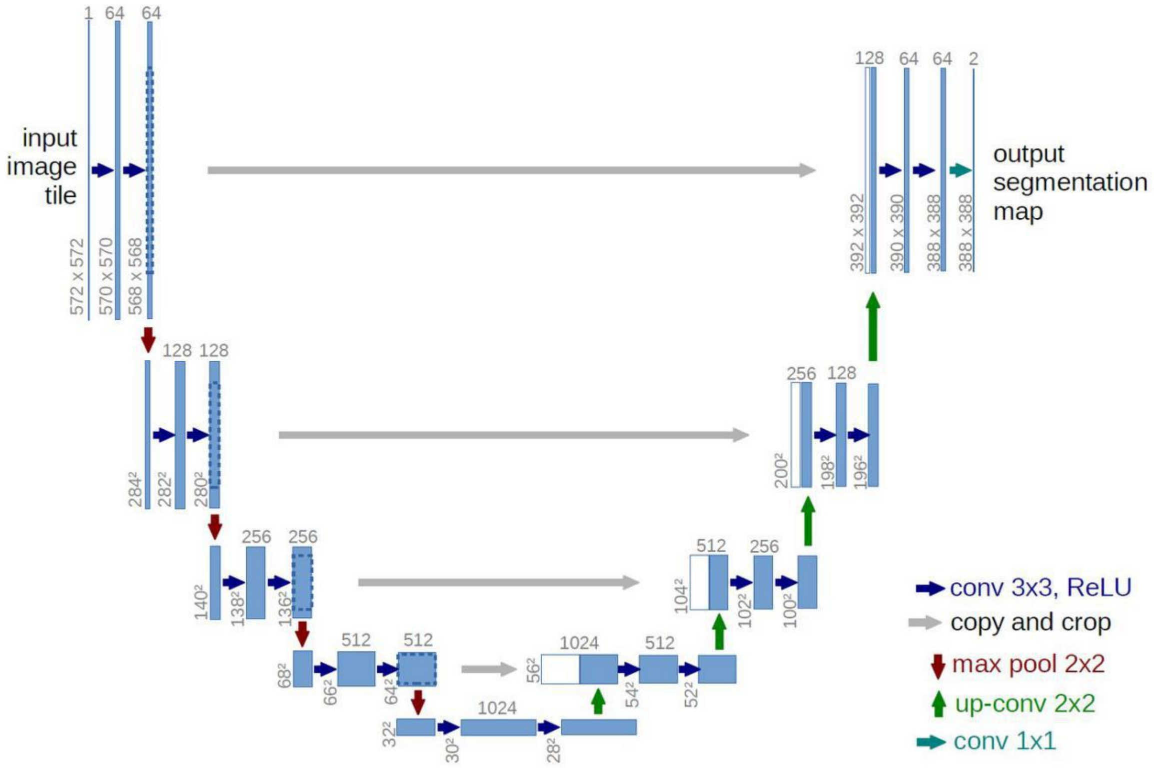

Figure 2: Flow chart of thyroid nodule segmentation (https://zhuanlan.zhihu.com/p/43927696)

Loss calculation in U_Net: In the U_Net method, in order to make the edge of the segmented object have a higher weight, a new loss weighting method needs to be used for each pixel. The U_Net model can then be used to segment biomedical images in a discontinuous manner for better identification of nodules in binary segmented images. The soft max function, also known as the normalized exponential function, is applied to reflect the result of the segmentation in the form of probability[6]. Then apply the cross-entropy loss function. In segmentation, we want to classify each pixel into a class, each pixel needs to exist in a certain class, just make sure they are in the class[7]. So, the segmentation problem is transformed into a multi-class classification problem, which works better than the traditional loss function.

The functional expression of cross entropy is as follows:

$$H\left(p,q\right)=-\sum_{i=1}^{n}p\left(x_i\right)\log\left(q\left(x_i\right)\right)$$

It can be seen that the formula of cross entropy is similar to relative entropy. Cross entropy is a variant of relative entropy. Since we have supervised training during model training, the sample label has been determined, so the real probability distribution P(x) is known, the expression of H(X) is as follows:

$$H\left(X\right)=-\sum_{i=1}^{n}p\left(x_i\right)\log\left(p\left(x_i\right)\right)$$

H(X) is a fixed value, approximately constant.，the relative entropy is shown as follows:

$$D_{KL}\left(p\,||\,q\right)=cons\tan t-\sum_{i=1}^{n}p\left(x_i\right)\log\left(q\left(x_i\right)\right)$$

When it is used as a loss function, the constant can be ignored, and the representation of cross entropy is obtained.

3 EVALUATION INDICATORS

In the field of computer vision deep learning image segmentation, mIoU value is an important indicator to measure the accuracy of image segmentation. as shown in the formula:

$$\mathrm{m}IoU=\frac{TP}{TP+FP+FN}$$

FN is the number of positive samples that were wrongly determined as negative samples, TN was the number of negative samples that were wrongly determined, FP was the number of negative samples that were wrongly determined to be positive samples , and TP was the number of positive samples that were correctly determined. If it is specific to the segmentation of thyroid nodules, positive samples are pixels in the nodule area, and negative samples are other pixels.

Pixel accuracy (PA), the simplest metric, is used to calculate the ratio between the number of correctly classified pixels and the total number of pixels.The functional expression of PA is as follows:

$$PA=\frac{TP+TN}{TP+TN+FP+FN}$$

The accuracy rate refers to the proportion of the samples that are actually positive among the samples that are predicted to be positive by the model.The functional expression of Precision is as follows:

$$\Pr ecision=\frac{TP}{TP+FP}$$

Recall rate refers to the proportion of the predicted positive samples among the actual positive samples.The functional expression of Recall is as follows:

$$\operatorname{Re}call=\frac{TP}{TP+FN}$$

F1 score is the harmonic mean of Precision and Recall , which is an indicator used to measure the accuracy of the two-class model.The functional expression of F1s is as follows:

$$F1\mathrm{s}=\frac{2*\Pr ecision*\operatorname{Re}call}{\Pr ecision+\operatorname{Re}call}$$

4 RESULT

In this paper, b_mIoU is used to represent the intersection ratio of the background, that is, the ratio of the area of the segmented background to the sum of the area of all backgrounds and nodules, and n_mIoU represents the intersection ratio of nodules , that is, the area of the segmented nodules and the total The ratio of the sum of background and nodule area, and the mean intersection ratio mIoU represents the average of b_mIoU and n_mIoU .The results of this experiment are shown in Table 1：

Table 1: Evaluation Metrics

Evaluation indicators	b_mIoU(%)	n_mIoU(%)	mIoU(%)	PA(%)	F1s(%)
Index value	95.16	61.31	78.24	95.5	86.8

Figure 3 is the segmentation result map, the first column of images is the thyroid ultrasound image, the second column of images is the label map, and the third column of images is the segmentation result obtained by this method.

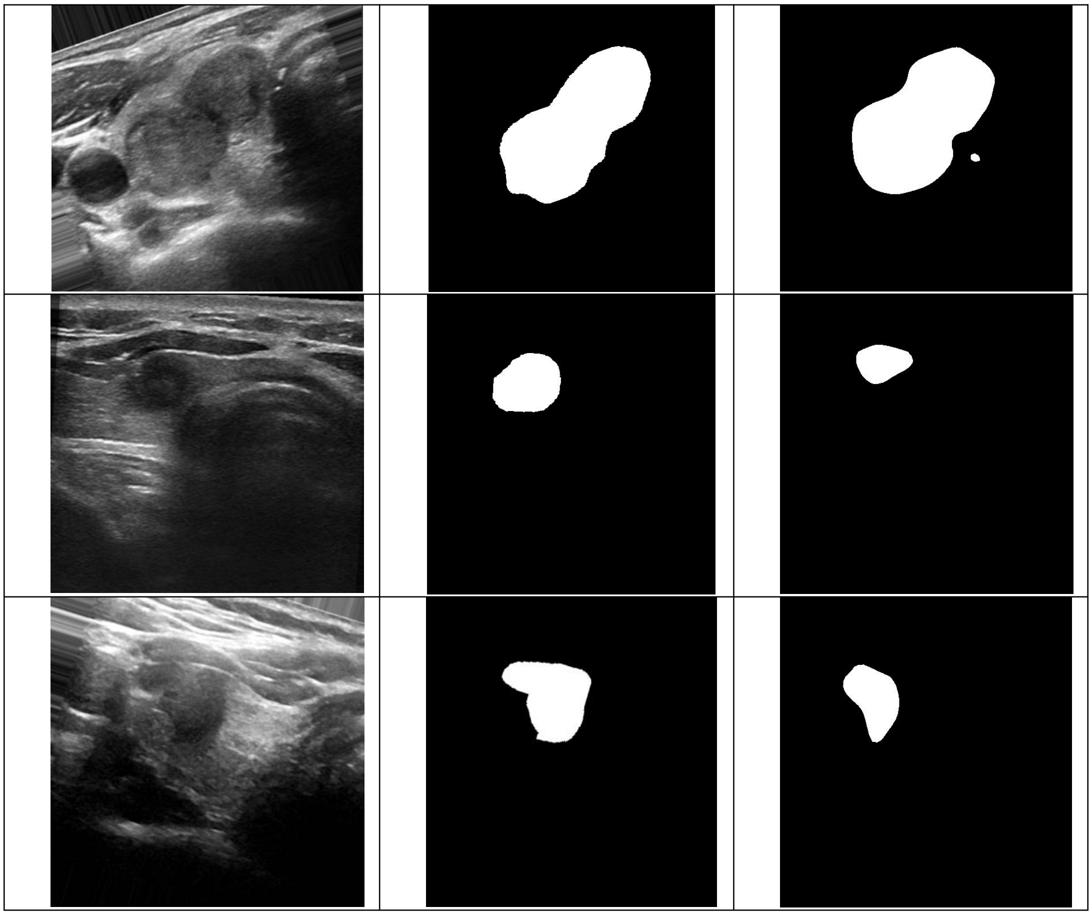

Figure 3 : Results of thyroid nodule segmentation

5 SUMMARY

Deep learning algorithms have great clinical value for the study of medical images, and can help doctors quickly and accurately determine the location, size, shape and other information of thyroid nodules, which is of great significance for doctors' diagnosis. A large number of high-quality datasets are very helpful for research, and it is hoped that more public datasets can be used by researchers. There is room for improvement in the segmentation method used in this paper, which can continue to improve the accuracy of segmentation on the existing basis.

6 COMPUTER CODE

```
def fast_hist(a, b, n):
    k = (a >= 0) & (a < n)
       return np.bincount(n * a[k].astype(int) + b[k], minlength=n ** 2).reshape(n, n)
def evaluation(hist, n_classes):
    TP = np.diag(hist)
    FP = hist.sum(0) - TP
    FN = hist.sum(1) - TP
    mIOU = TP / (TP + FP + FN)
    Precision = TP / (TP + FP)
    Recall = TP / (TP + FN)
    PA = TP.sum(0) / (TP.sum(0) + FP.sum(0))
    F1 = (2 * Precision * Recall) / (Precision + Recall)
    mF1 = F1.sum(0) / n_classes
    return mIOU, PA, mF1
def SegColor2Label(img):
    VOC_COLORMAP = [[0, 0, 0], [255, 255, 255]]
    mapMatrix = np.zeros(256 * 256 * 256, dtype=np.int32)
    for i, cm in enumerate(VOC_COLORMAP):
        mapMatrix[cm[2] * 65536 + cm[1] * 256 + cm[0]] = i
    indices = img[:, :, 0] * 65536 + img[:, :, 1] * 256 + img[:, :, 2]
    return mapMatrix[indices]
def Evaluation(test_label_dir, pred_dir, name_index_map, n_classes):
    hist = np.zeros((n_classes, n_classes))
    label_path_lists = os.listdir(test_label_dir)
    for i, label_path_list in enumerate(label_path_lists):
        label = cv2.imread(os.path.join(test_label_dir, label_path_list))
```

```
        label = cv2.resize(label, (512, 512))
        label[label != 0] = 1
               pred = cv2.imread(os.
                          path.join(pred_dir, label_path_list))
        pred[pred != 0] = 1
        if len(label.flatten()) != len(pred.flatten()):

print('Skipping:len(gt)={:d},len(pred)={:d},{:s},{:s}'.format(len(label.flatten()),
len(pred.flatten()),
os.path.join(test_label_dir,
label_path_list),
os.path.join(pred_dir,
label_path_list)))
            continue
        hist += fast_hist(label.flatten(), pred.flatten(), n_classes)
        if i > 0 and i % 10 == 0:
            mIOU, PA, F1 = evaluation(hist, n_classes)
            print('第{:d}步  mIOU：{}  PA：{}  F1：{}'.format(i, mIOU, PA, F1))
    mIoUs, PAs, F1s = evaluation(hist, n_classes)
    for ind_class in range(n_classes):
    print('===>' + name_index_map[ind_class] + ':\t' + str(round(mIoUs[ind_class] * 100,
2)))
    print('===> mIoU: ' + str(round(np.nanmean(mIoUs) * 100, 2)))
    print('===> PA: {}'.format(PAs))
    print('===> F1s: {}'.format(F1s))
    return mIoUs, PAs, F1s
if __name__ == "__main__":
    test_label_dir = './data/test/labels'
    pred_dir = './Results/U_Net/pre'
    name_index_map = {0: 'b_mIoU', 1: 'n_mIoU'}
    Evaluation(test_label_dir, pred_dir, name_index_map, 2)
```

REFERENCES

[1] 高丹,陈建英,谢盈.A-PSPNet:一种融合注意力机制的 PSPNet 图像语义分割模型[J].中国电子科学研究院学报,2020,15(06):518-523.

[2] Bosch A, Zisserman A, Munoz X. Image classi， fication using random forests and ferns.[C].Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on. EEE, 2007: 1-8.

[3] Long J, Shelhamer E, Darrell T. Fully convolutional networks for smantic segmentation.[C]. EEE Conference on Computer Vision and Pattern Recognition. EEE Computer Society, 2015:3431-3440.

[4] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]. Intemational Conference on Medical Image Computing and Computer-. Assisted Intervention. Springer, Cham, 2015:234-241.

[5] Ma J, Wu F, Zhao Q, et al. Ultrasound image-based thyroid nodule automatic segmentation using convolutional neural networks[J]. Intemational journal of computer assisted radiology and surgery, 2017, 12(11): 1895-1910.

[6] 贾树开. 深度学习在图像分割中的应用[D].电子科技大学,2020.DOI:10.27005/d.cnki.gdzku.2020.004196.

[7] 王小燕,江建伟,徐龙滔.基于 CMCP 和余弦间隔交叉熵的深度神经网络及其应用[J/OL].数量经济技术经济研究:1-19[2022-09-22].DOI:10.13653/j.cnki.jqte.20220920.006.

***This form below helps us to understand your paper better, so please fill in the information of all authors. The form itself will not be published.**

Authors’ background

Position can be chosen from:

Postgraduate / Ms.

Paper ID	Position , Full Name, Working unit & nation	Email address	Research Interests	Personal website (if any)
C001	Postgraduate. Shandong Qingdao, Zhou Xueting, Shandong University of Science and Technology, China	202182130042@sdust.edu.cn	image processing	