Automatic cardiac segmentation methods for MS-CMR images

Yuxin Li
SDS, Fudan University, Shanghai, China

Abstract

Accurate assessment of myocardial viability in multi-sequence cardiac magnetic resonance(CMR) images is desired to automate disease diagnosis. To classify myocardial pathology automatic segmentation methods are necessary. In this paper, we propose to use an automatic segmentation for each slice with convolutional neural network architecture based on U-net. We compare the performances of different networks to segment the CMR images. A five-fold cross-validation strategy is utilized to assess the performance of the proposed methods. Our models were evaluated by the 2019 Multi-sequence Cardiac MR Segmentation Challenge. The experiment result shows a satisfying performance of the proposed framework. Code is available at https://github.com/Grit1021/U-net-seg.

Keywords: Cardiac MR; Image segmentation; Convolutional neural networks

1 Introduction

Cardiac MRI is a significant technology for cardiac function analysis. Benefiting from this technology, the doctor can evaluate the heart function non-invasively. Analysis of myocardial(Myo) viability is crucial to better understand the physiological and pathological processes of patients suffering from myocardial infraction(MI). For quantitative assessment, the segmentation of the myocardium is a prerequisite.However, manual segmentation can be time-consuming and suffer from inter-observer variations, thus automating this process is desirable in the clinic.

The segmentation of cardiac MRI data plays an important role in the process of diagnosing the disease, which has been a key challenge, and machine learning techniques have heavily influenced the trends. Earlier approaches are developed based on semi-automated and atlas-based to perform the segmentation of cardiac MR images Poudel, Lamata, and Montana (2017).In recent years, many semi-automated and automated methods have been proposed for multi-modal medical image segmentation using deep learning-based methods, such as convolutional neural networks (CNNs) LeCun, Bottou, Bengio, and Haffner (1998) and fully convolutional networks Long, Shelhamer, and Darrell (2015) especially the U-Net architecture Ronneberger, Fischer, and Brox (2015). Specifically, the U-shape network structure plays a significant role in medical image segmentation, and numerous variants have been developed. Chen et al. Chen et al. (2018) proposed a two-task recursive attention model to segment the left atrium and the atrial scars simultaneously. Wang et al. Wang et al. (2020) presented SK-Unet to segment the left ventricle, right ventricle and left ventricular myocardium, which embeds attention modules in different stages of the structure based on U-net. Furthermore, there are also several works aiming at the segmentation of cardiac anatomical structures in CMR Yue, Luo, Ye, Xu, and Zhuang (2019).

In this paper, we utilize different networks for cardiac pathology segmentation and compare the performances. The architectures have not been the only focus for network performance improvement in cardiac MRI segmentation. Several loss functions, such as weighted cross-entropy Jang, Hong, Ha, Kim, and Chang (2018), Dice loss Milletari, Navab, and Ahmadi (2016), and weighted Dice loss Yang, Bian, Yu, Ni, and Heng (2018) have been investigated to increase the alignment with ground truth annotations. The problem of regional imbalance can be alleviated by advancing the loss functions. In recent years, multiple challenges on the segmentation task have also been opened for advancing the state of the art for segmentation tasks Bernard et al. (2018).In this work, we showcase the performance of several networks based on CNN and U-net architectures on myocardial pathology segmentation.

The rest of this article is organized as follows: we introduce our method in Section 2. Results are analyzed in Section 3. Finally, we conclude this paper in Section 4.

2 Method

Our proposed model is based on the basic convolutional neural network and classical U-net, which are effective in medical image segmentation tasks.

2.1 Dataset Description and Pre-processing

Dataset description The dataset is coming from the 2019 Multi-sequence Cardiac MR Segmentation Challenge (MS-CMR2019) Qiu et al. (2023). It published 45 patients' CMR data with three modalities having a different number of slices of multi-sequence CMR,i.e., T2,balanced-Steady State Free Precession(b-SSFP), and late gadolinium enhancement(LGE), and all ground truth values for every single slice. There are 35 labeled T2 CMR data with about 3 slices of each patient, 35 labeled b-SSFP CMR data with about 11 slices of each patient, and just 5 labeled CMR data with about 15 slices of each patient. The rest volumes are unlabeled data. The main purpose of this challenge is to segment the left ventricle (LV), right ventricle (RV) and left ventricle myocardium (LVM) from LGE CMR data.Correspondingly, the ground truth labels include left ventricular (LV), right ventricular (RV), and left ventricle myocardium (LVM). However, rarely labeled target data increases the challenge sharply.

The following is an example of a multi-sequence CMR image from the training dataset. First image bSSFP, second image LGE, third image T2 CMR, and their corresponding images associated ground truth segmentation. The cardiac region is segmented into LV, RV, and LVM regions, where dark green represents the right ventricle, red represents the left ventricular, and white represents the left ventricle myocardium.

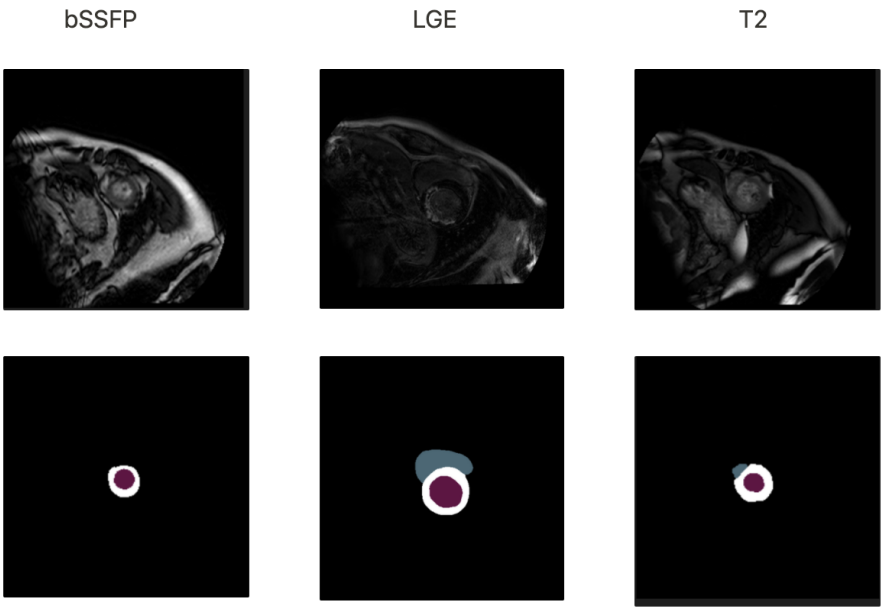

Figure 1: An example of the dataset

Data augmentation In case of insufficient data, it is necessary to increase the number of images by providing diversity in the dataset in order to prevent over-fitting. Data augmentation techniques include image 45° to 315° rotation, horizontally flipping $\frac{2}{3}$ of the images, elastic transformations, and random scaling. The data augmentation method was applied to both original images and ground truth masks.

2.2 Networks architectures

FCN To effectively utilize the end-to-end style of deep learning, Long et al. (2015) converted CNN for image-wise classification to FCN that aims at pixel-wise semantic segmentation. In the framework of FCN, FC layers are replaced by a convolution layer with a 1×1 size kernel. To recover the size of feature maps, deconvolution layers that most commonly adopt bilinear interpolation are used to reverse the downsampling process. In fact, FCN for semantic image segmentation is a particular case of CNN, which is trained from end to end and pixel to pixel. FCN with a fusion strategy computes score maps among per-class probability maps interpolated from different-level units.

Segnet Segnet is a novel and practical deep fully convolutional neural network architecture for semantic pixel-wise segmentation. This core trainable segmentation engine consists of an encoder network, and a corresponding decoder network followed by a pixel-wise classification layer. The architecture of the encoder network is topologically identical to the 13 convolutional layers in the VGG16 network Simonyan and Zisserman (2014). The role of the decoder network is to map the low-resolution encoder feature maps to full input-resolution feature maps for pixel-wise classification. The novelty of SegNet lies is in the manner in which the decoder upsamples its

lower-resolution input feature map(s). Specifically, the decoder uses pooling indices computed in the max-pooling step of the corresponding encoder to perform non-linear upsampling. This eliminates the need for learning to upsample. The upsampled maps are sparse and are then convolved with trainable filters to produce dense feature maps.

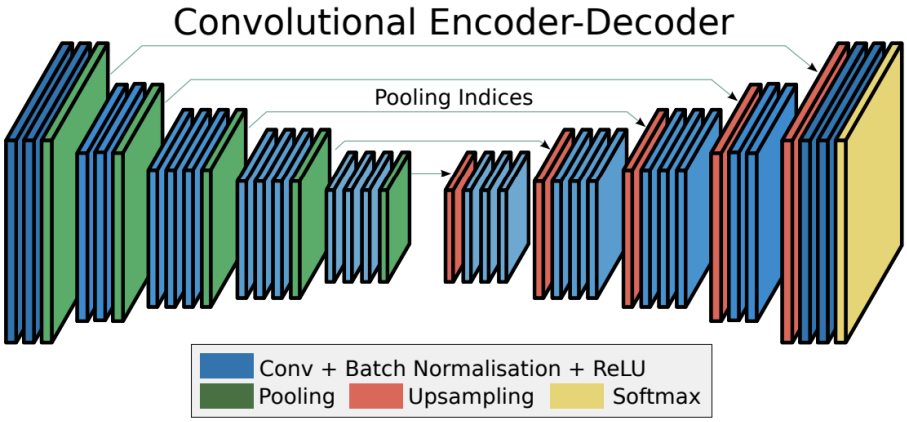

Figure 2: Segnet structure

U-net The U-Net architecture, a convolutional network for biomedical image segmentation is used to segmentation of the cardiac MR images. The network consists of contracting (downsampling) and expanding (upsampling) paths. In the contracting path, there are 3×3 convolutions, each followed by a rectified linear unit (ReLU) activation function and 2×2 max-pooling operation with stride 2. There is also a dropout layer after the first convolution at each downsampling step to prevent the architecture from overfitting. In the contracting step the number of feature channels doubles. In the expansive, there is upsampling of the feature map followed by a 2×2 convolution in all steps. At each step, the number of feature channels halves. There is also a concatenation with the feature map from the contracting path, and two 3×3 convolutions, each followed by a ReLU activation function. A 1×1 convolution also is used to map each feature vector to the number of classes at the final layer.

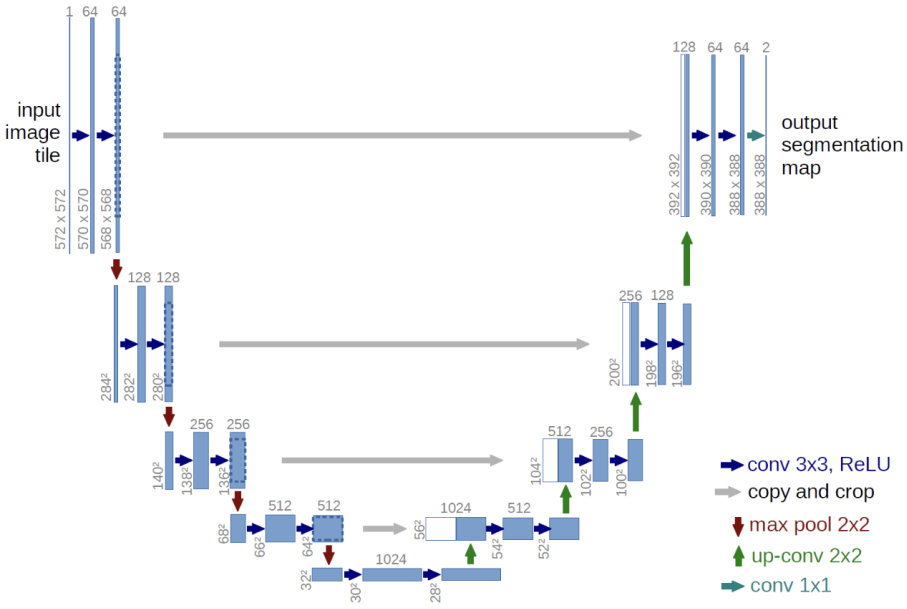

Figure 3: U-net structure

attU-net The architecture of the sub-network contains attention blocks in the original U-net and we call it attention-Unet. Different from the original Unet, attention-Unet has attention blocks in the network. Attention blocks in the encoder are called feature attention blocks, and those in the decoder which take the place of the

concatenation part in the original Unet are called concate attention blocks. Apart from these, we apply group convolutions in attention-Unet for model compression.

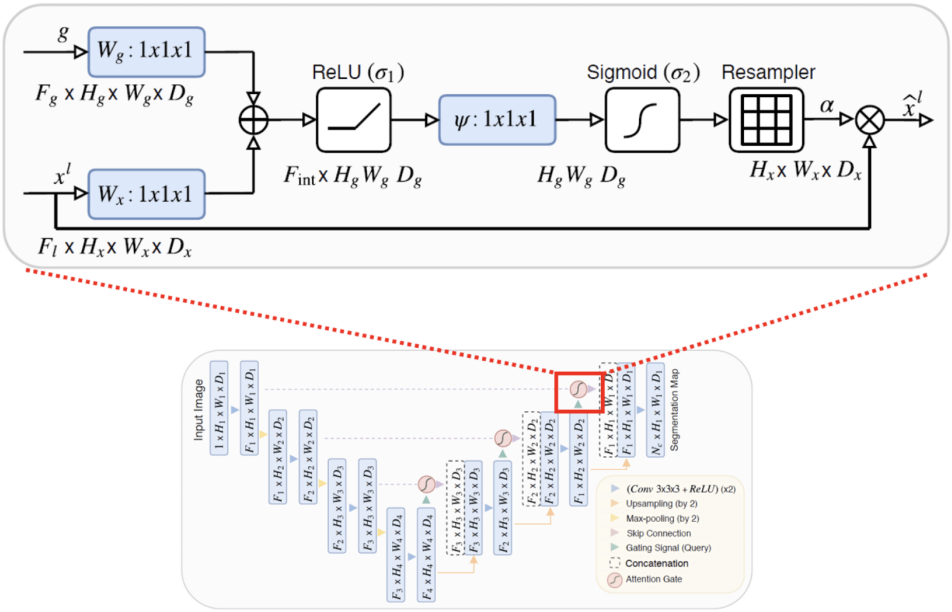

Figure 4: attU-net structure

2.3 Training implementation

In order to maximize the data utilization, we train FCN, Segnet, U-net, and attU-net architectures using 5-fold cross-validation in the training set. Moreover, we also define early stopping according to our dataset. Each model has trained 100 epochs with optional learning rates and 36 batch sizes. The training time is about 1 hour for each model. We adopt adam optimizer and use different kinds of loss functions. Notably, we utilize a weighted cross entropy (wbce) loss function to solve the class imbalance problem.

Our segmentation model is evaluated by the official evaluation metrics, which are the Dice score, Jaccard score, and Hausdorff distance. Dice score and Jaccard score are overlapped metrics. They evaluate the overlap ratio between the ground truth and the predicted result. However, they have a shortage of the boundary details of the subject. Although similarity metrics, Hausdorff distance, is mainly focused on the similarity between the ground truth and predicted result, it is sensitive to noise. Utilizing both of these metrics can complement one another perfectly. Hence, the segmentation model can be over-all evaluated. Notice that the Dice score is the main metric.

3 Experimental results

The score of metrics during the validation stage is shown in Table 1 and the numbers in bold are the best values of the four methods. Without loss of generality, we list the results of the T2 CMR images in fold 5. These scores are the mean value of all the slices, which are calculated by average operation without weight. The segmentation model has a satisfying performance on the overlap metrics. Due to the model being trained on the short axis, the performance on the similarity metrics is worse than on overlap metrics.

The training process requires a loss function to update the model parameters through backpropagation to minimize the loss function. The architectures of the following table are trained with dice loss which is defined as follows:

$$L(y, \hat{y}) = -\frac{1}{M}\sum_{j=0}^{M}\sum_{i=0}^{N}(y_{ij} log(\hat{y}_{ij})) \tag{1}$$

where $\hat{y}$ is the predicted expected value, y is the observed value for each class N. Dice coefficient used for segmentation accuracy assessment by evaluating the overlap between the ground truth and the predicted area and defined as:

$$DC = \frac{2|A \cap B|}{|A| + |B|} \tag{2}$$

where A refers to manual segmentation and B refers to automated segmentation area. The $1 - DC$ is called soft-Dice loss function Milletari et al. (2016).

Table 1: Experiment result table

Network	Evaluation category	Mean	LVM	LV	RV
FCN	Dice score	0.86±0.095	0.895±0.0356	0.914±0.057	0.771±0.243
	Jaccard score	$\mathbf{0.777 \pm 0.104}$	$\mathbf{0.812 \pm 0.0559}$	0.846±0.088	$\mathbf{0.673 \pm 0.235}$
	Hausdorff distance	7.14±1.73	7.46±3.28	3.79±1.31	10.2±4.18
Segnet	Dice score	0.862±0.105	0.896±0.042	0.912±0.0667	0.776±0.251
	Jaccard score	0.781±0.122	0.814±0.0652	0.845±0.101	0.685±0.249
	Hausdorff distance	7.37±1.98	7.74±3.01	4.23±1.77	10.1±3.59
U-net	Dice score	$\mathbf{0.883 \pm 0.0964}$	$\mathbf{0.914 \pm 0.0388}$	$\mathbf{0.935 \pm 0.0495}$	$\mathbf{0.8 \pm 0.249}$
	Jaccard score	$\mathbf{0.814 \pm 0.107}$	$\mathbf{0.845 \pm 0.0623}$	$\mathbf{0.881 \pm 0.0791}$	$\mathbf{0.717 \pm 0.246}$
	Hausdorff distance	6.76±1.69	6.87±1.92	3.57±1.34	9.86±4.39
attU-net	Dice score	0.877±0.0982	0.913±0.0389	0.93±0.0462	0.788±0.249
	Jaccard score	0.805±0.11	0.843±0.0619	0.873±0.0753	0.7±0.246
	Hausdorff distance	$\mathbf{6.66 \pm 1.81}$	$\mathbf{6.74 \pm 1.64}$	$\mathbf{3.65 \pm 1.02}$	$\mathbf{9.6 \pm 4.7}$

From the table above, we notice that the performances among these network structures are close to one another. However, due to the simplicity of FCN and Segnet, they consume less time. U-net has better performance than attU-net, which certifies the limited effect of attention block. We propose U-net as the best model due to its efficiency.

An example of CMR image segmentation validation result in T2 is the 5th fold of different methods is shown in the following figures. We select a representative slice to show the result with different methods. The result shows that our prediction contours can perfectly fit the ground truth.

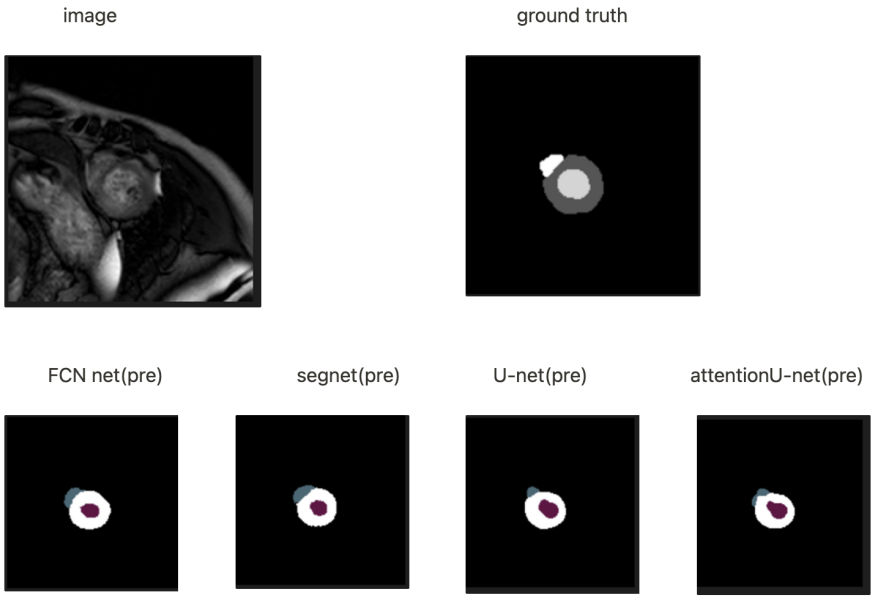

Figure 5: segmentation result

4 Discussion and Conclusion

CMR is an efficient technology to identify infarcted myocardium. In this paper, we proposed an automatic framework for CMR segmentation and compared several different methods' performances. The model based on a

deep convolutional network mostly outperformed the performance of U-Net. The final segmentation is evaluated by the organizer, and the mean metrics score of Dice score, Jaccard score, and Hausdorff distance are utilized. Some test cases have been challenging to segment for the proposed technique due to inconsistent intensity patterns in the edges. One avenue of improvement could be to incorporate temporal information from multiple temporal slices. Further investigation will be devoted to alleviating the problem of the efficiency of the network iterations.

References

Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P.-A., ... others (2018). Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE transactions on medical imaging*, *37*(11), 2514–2525.

Chen, J., Yang, G., Gao, Z., Ni, H., Angelini, E., Mohiaddin, R., ... others (2018). Multiview two-task recursive attention model for left atrium and atrial scars segmentation. In *International conference on medical image computing and computer-assisted intervention* (pp. 455–463).

Jang, Y., Hong, Y., Ha, S., Kim, S., & Chang, H.-J. (2018). Automatic segmentation of lv and rv in cardiac mri. In *Statistical atlases and computational models of the heart. acdc and mmwhs challenges: 8th international workshop, stacom 2017, held in conjunction with miccai 2017, quebec city, canada, september 10-14, 2017, revised selected papers 8* (pp. 161–169).

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324.

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 3431–3440).

Milletari, F., Navab, N., & Ahmadi, S.-A. (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3d vision (3dv)* (pp. 565–571).

Poudel, R. P., Lamata, P., & Montana, G. (2017). Recurrent fully convolutional neural networks for multi-slice mri cardiac segmentation. In *Reconstruction, segmentation, and analysis of medical images: First international workshops, rambo 2016 and hvsmr 2016, held in conjunction with miccai 2016, athens, greece, october 17, 2016, revised selected papers 1* (pp. 83–94).

Qiu, J., Li, L., Wang, S., Zhang, K., Chen, Y., Yang, S., & Zhuang, X. (2023). Myops-net: Myocardial pathology segmentation with flexible combination of multi-sequence cmr images. *Medical Image Analysis*, *84*, 102694.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International conference on medical image computing and computer-assisted intervention* (pp. 234–241).

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Wang, X., Yang, S., Tang, M., Wei, Y., Han, X., He, L., & Zhang, J. (2020). Sk-unet: an improved u-net model with selective kernel for the segmentation of multi-sequence cardiac mr. In *Statistical atlases and computational models of the heart. multi-sequence cmr segmentation, crt-epiggy and lv full quantification challenges: 10th international workshop, stacom 2019, held in conjunction with miccai 2019, shenzhen, china, october 13, 2019, revised selected papers 10* (pp. 246–253).

Yang, X., Bian, C., Yu, L., Ni, D., & Heng, P.-A. (2018). Class-balanced deep neural network for automatic ventricular structure segmentation. In *Statistical atlases and computational models of the heart. acdc and mmwhs challenges: 8th international workshop, stacom 2017, held in conjunction with miccai 2017, quebec city, canada, september 10-14, 2017, revised selected papers 8* (pp. 152–160).

Yue, Q., Luo, X., Ye, Q., Xu, L., & Zhuang, X. (2019). Cardiac segmentation from lge mri using deep neural network incorporating shape and spatial priors. In *Medical image computing and computer assisted intervention–miccai 2019: 22nd international conference, shenzhen, china, october 13–17, 2019, proceedings, part ii 22* (pp. 559–567).

Zhuang, X. (2019). Multivariate mixture model for myocardial segmentation combining multi-source images. *IEEE transactions on pattern analysis and machine intelligence, 41* (12), 2933-2946

Supplementary Files (optional)

The source code can be accessed via the following URL:https://github.com/Grit1021/U-net-seg.