Recurrent residual U-Net for medical image segmentation

P.Vamsi Kumar
vamsi81523@gmail.com

Abstract

Since it's inception in 2015 U-Net has become one of the most popular choice for Medical Image Segmentation tasks. The base U-Net which was proposed in 2015 contained convolutional units. Authors of the R2U-Net model [1] have introduced Recurrent and Residual Networks in the base convolutional units.The proposed model has been tested on three kind of datasets such as blood vessel segmentation in retinal images, skin cancer segmentation and lung lesion segmentation.

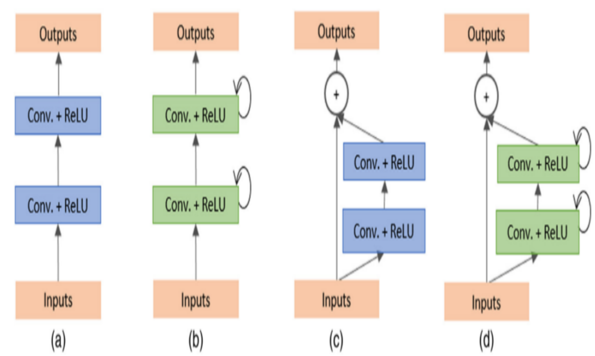

Figure 1. Different variants of the convolutional and recurrent convolutional units (RCUs) including (a) the forward convolutional unit,(b) the recurrent convolutional unit, (c) the residual convolutional unit, and (d) the recurrent residual convolutional unit.

1. Introduction

Medical Image Segmentation is a very tedious task to do manually and labeling the dataset requires an expert in this field which is expensive and it requires a lot of effort and time due to this very reason most of the time a good number of labels are not available for training.There were several traditional machine-learning and image-processing techniques available for medical image segmentation task before the rise of Deep Learning. But segmentation approaches that utilize DL have become very popular in recent years. And particularly U-Net model has been performing very well on the medical image segmentation task of different modalities. The kind of convolutional unit which is used in base U-Net can be seen in the Fig:1(a) this doesn't consist of any Residual or Recurrent networks and the convolutional unit which is presented in Fig:1(d) is the one which authors have used in the R2U-Net model.

Architecture for R2U-Net has been shown in Fig:2. In the encoding part each convolutional unit consists of convolutional layers, residual connection and recurrent layer as shown in Fig:1(d) followed by ReLU, max-pooling layer. In the decoding part each convolutional unit consists of convolutional layers, residual connection and recurrent layer followed by ReLU and transpose convolution layer for upsampling.

2. Datasets

The proposed model has been tested on three different medical imaging datasets like blood vessel segmentation from retina images (DRIVE,STARE,CHASE_DB1),skin cancer lesion segmentation and lung segmentation.

2.1. Blood Vessel from Retina Images Dataset

The DRIVE dataset is consisted of 20 color retinal images and size of each original image is 565×584 pixels. STARE dataset consists of 20 color images and each image is of size 700×605 pixels. CHASE dataset consists of 28 color retina images and each image is of size 999×960 pixels.Patch based approach is used for training the model on these datasets. For each dataset 1,00,000 patches of size 48×48 are created randomly.And 80% of the input images are used for training and 10% are used for validation and the other 10% are used for testing.

2.2. Skin Cancer Segmentation

This dataset is taken from Kaggle competition on skin lesion segmentation that occured in 2016.The dataset consists of 900 images and each image is of size 700×900 which is rescaled to 128×128. For train:val:test ratio 80:10:10 is

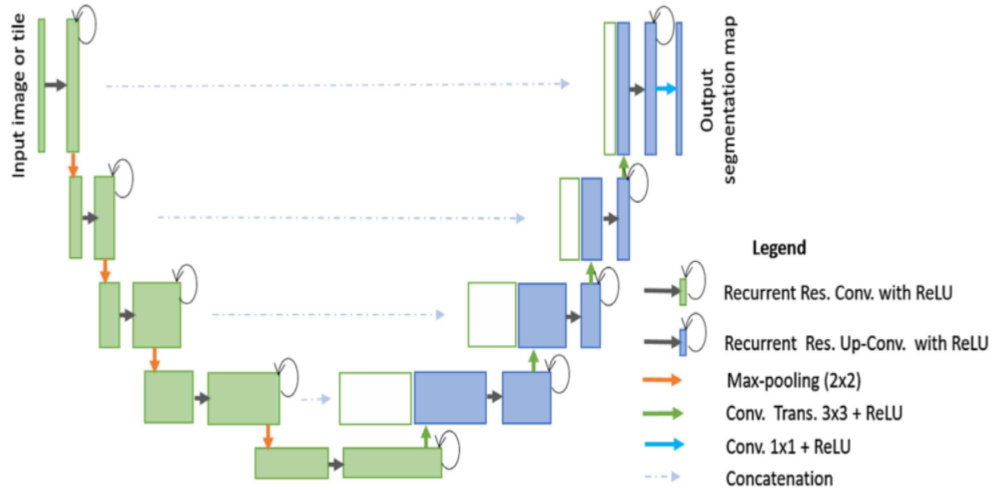

Figure 2. The R2U-Net architecture with convolutional encoding and decoding units using RCLs, which is based on a U-Net architecture.

applied.

2.3. Lung Segmentation

This dataset is taken from LUNA-16 competition at Kaggle Data Science Bowl in 2017. It consists of 267 2-D samples each image of size 512×512 which has been rescaled to 256×256 during the implementation.

3. Methodology

Each convolutional unit in the final model consists of two convolutional layers along with residual and recurrent networks as shown in Fig:1(d) and after that ReLU is used as activation function and for downsampling 2×2 max-pooling operations are performed. In the decoding phase the convolution transpose operations are performed to up-sample feature maps. Features from encoding units are concated to the features in decoding units. And finally 1×1 convolutional kernels are used with sigmoid function at the output.The re-implemented model consists of three encoding and three decoding units.And Binary Cross Entropy loss is used for training the model.And Accuracy is used as the metric for testing purposes.Adam optimization technique is used with various learning rates for different datasets. As there are multiple datasets different batch sizes are used for these datasets and for this purpose a DataLoader class has been created in PyTorch for each dataset.And each dataset has been trained for 100 epochs.

4. Results

The results from the re-implemented model are very similar to the ones that are published in the paper.But for the Retinal Image Datasets authors have used created around 2,50,000 patches while training the model with these number of inputs each epoch was taking around 30 minutes so the number of patches have been reduced from this original number this might be a factor behind the difference in accuracies.

5. Conclusion

There is an improvement in the results from U-Net to R2U-Net model.The re-implemented model have two convolutional layers in each encoding and decoding unit.Authors have tested by increasing convolutional layers from 2 to 3 and they were getting slight improvement.Recently a paper on Attention U-Net model has been published and it would be interesting to see if we can get any improvement by adding Attention into the R2U-Net model.

Dataset	From Paper	Re-Implemented Results
DRIVE	0.955	0.951
STARE	0.971	0.962
CHASE	0.963	0.957
Skin Lesion	0.947	0.939
Lung Segmentation	0.994	0.991

Table 1. Validation Accuracy Comparison

References

[1] M.Z.Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, "Recurrent residual U-Net for medical image segmentation," J. Med. Imag., vol. 6,no. 1, Mar. 2019, Art. no. 014006.