W-net: Bridged U-net for 2D Medical Image Segmentation

Wanli Chen[a,1], Yue Zhang[a,b,1], Yifan Chen[c], Hongjian Shi[a,*], Xiaoying Tang[a,*]

[a]*Department of Electrical and Electronic Engineering, Southern University of Science and Technology, China*
[b]*Department of Electrical and Electronic Engineering, The University of Hong Kong, China*
[c]*Faculty of Science and Engineering, The University of Waikato, New Zealand*

Abstract

This document briefly describes techniques we used in automatic segmentation of the prostate in transversal T2 MRI for the PROMISE12 challenge. We used bridging, skip connection, ReLU and ELU cluster and our proposed cos-diceloss to enhanced the performance of stacked U-nets, namely W-net.

1. Cos-Dice Loss

We propose Cos-Dice Loss Function:

$$L_{CosDice} = \cos^{Q}\left(\frac{\pi}{2} \cdot DSC\right), Q > 1. \tag{1}$$

Where Q is an adjustable number. As it shown in Fig. 1, the cos-dice loss is smoother than dice loss when the intersection percentage is large and rougher than dice loss when the intersection percentage is small.

2. Network architecture

Our network is based on U-net[1], which is a classical encoder-decoder net in medical image application. Based on U-net, a stacked U-net is proposed.

*Corresponding author
Email addresses: `shihj@sustc.edu.cn` (Hongjian Shi), `tangxy@sustc.edu.cn` (Xiaoying Tang)
[1]Equal contribution

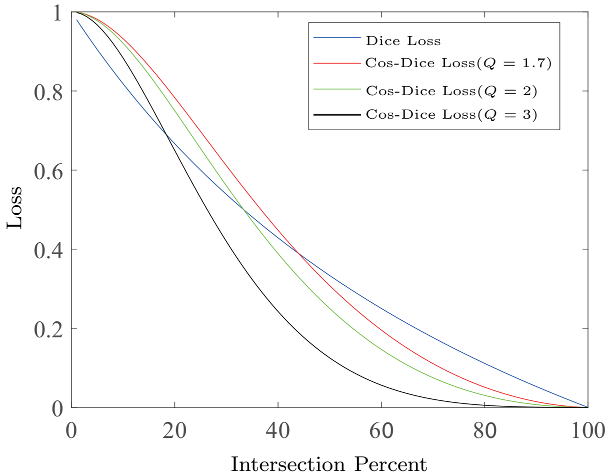

Figure 1: Dice loss and cos-dice loss with different factor Q

The stacked U-net improves network performance by using the first U-net to find a coarse feature and use the second U-net to obtain a fine result. The stacked U-net is, however, not useful for medical image segmentation. It is hard to reach convergence and usually dive into a sub-optimal solution because the increasing complexity of network. To overcome the issue, we propose a network bridging method. Different from the previous stacked U-net which acquires large number training data, bridging two U-nets can reduce the training cost and makes the network fit for medical application where the training data are usually not sufficient. This is because bridging two U-nets can fully use different features in multi levels, which will accelerate the convergence of neural network. Our network structure is shown on Fig. 2. The gray block represents a ELU cluster (2 conv-BN-ELU blocks), and the yellow block represents a ReLU cluster (2 conv-BN-ReLU blocks). The dotted lines represents network bridging. The red lines represents skip connections.

References

[1] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical

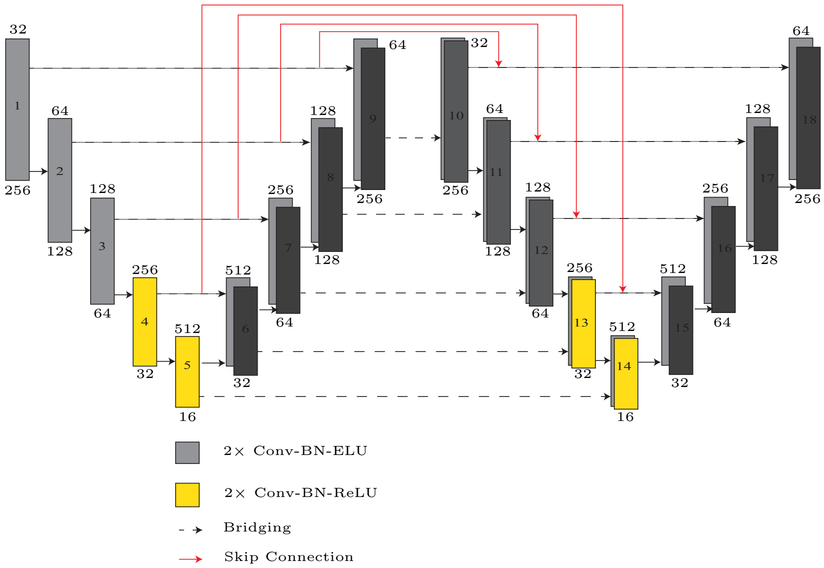

Figure 2: W-net architecture. The number above each block represents the number of feature channels. The number inside each block represents the sequence number. The number below each block means the image size.

image computing and computer-assisted intervention, Springer, 2015, pp. 234–241.