1. Literature Survey (2)

I analyzed the essay *IDRiD Diabetic Retinopathy Segmentation and Grading Challenge* [1] which is a comprehensive report of the Retinopathy Segmentation challenge. The manners of segmentation in this essay is mainly based on deep learning. We compared different neural network structures. Most of the structures used U-net and ensembled other structure. The team VRT in the challenge used original U-net structure and averagely got the best performance. Also, according to their essay[2], after reducing image resolution by about 6 times, compared to original resolution, U-net achieved less false positive and better sensitivity with lower resolution input. Using lower resolution image also alleviates computation pressure. Therefore, we referred their U-net structure which is elaborated in their essay.

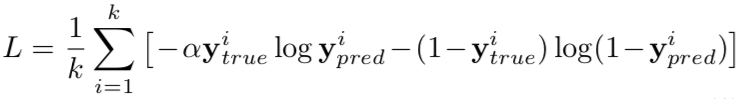
1. Method (4)
2. Image pre-processing:

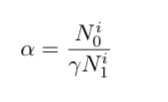
Firstly, we cropped original image from 4288 x 2848 to 3500 x 2848 by reducing its width. Secondly, we padding the image to 3500 x 3500 and followed by resizing it to 640 x 640 through bicubic interpolation to protect image details. Thirdly, to do normalization, each image is divided by 255 as the network input. Also, we did augmentation such as flipping, scaling, rotating and adding Gaussian noise.

1. Network Model And Loss Function:

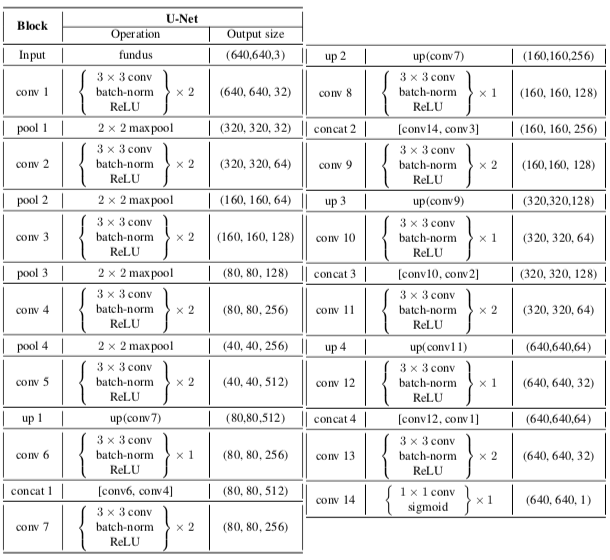
We used U-net[3] structure which is modified by team VRT. The input fundus image size is 640 x 640 x 3, the detail of convolutional blocks are shown as figure. After 4 times max-pooling, the image size is down to 40 x 40, and then the image is up-sampled to the original size with one channel. In the process of up-sampling, we concatenate the up-sampling layers with corresponding initial layers because the features are important in both initial layers and up-sampling layers for segmentation.

Because of the last layer of the model is sigmoid, the output of the image is in the range of [0, 1]. We used weighted binary cross entropy as loss function. In the function, k is the batch size. α is the weight which is calculated by. N0 is the number of background and N1 is the number of foreground. γ is the hyperparameter that we need to choose. The reason why we use the weight in the binary cross entropy is because the background area is significantly large in the image. We need to use the weight to reduce the false negative to reduce the loss.





1. Experiment:



1. Experimental Setup (4)
2. Results (3)
3. Discussion and Conclusions (2)
4. Reference:

[1] IDRiD\_Diabetic Retinopathy\_Segmentation and Grading Challenge

[2] ISBI 2018 SUB-CHALLENGE 1: LESION SEGMENTATION

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, “U-net: Convolutional networks for biomedical image segmentation,” in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015, pp. 234–241.

