ISIC 2018 – Skin Lesion Analysis

Chengyao Qian *[1], Hao Jiang†[1] and Ting Liu‡[1]

[1]MTLab, Meitu, Inc

July 28, 2018

Abstract

This paper summarises our method and validation results for the ISIC Challenge 2018 - Skin Lesion Analysis Towards Melanoma Detection - Task 1: Lesion Segmentation

1 Introduction

International Skin Imaging Collaboration (ISIC) provide various images of skin problems for early diagnosis. In this manuscript, we propose a method to address the Lesion Segmentation in the ISIC 2018. Our method is based on deep convolutional networks and trained by the dataset provided by ISIC [1] [2].

2 Materials and Methods

2.1 Database

For Lesion Segmentation, ISIC 2018 provides 2594 images with corresponding ground truth for training. There are about 100 images for validation and 1000 images for testing, the ground truth for validation and testing images are not provided. The images have different sizes from hundred to thousand, and the ratio of width and height is various. The lesion in images has different appearances and is located in different part of people. There are three criterions for ground truth. The whole training set is split into two sets with ration 10:1 for training and validation separately.

*qcy1@meitu.com
†jh1@meitu.com
‡lt@meitu.com

2.2 Methods

We design a process which includes detection and segmentation. Figure 1 shows our process. First, we use detection to find the bounding box of lesions in images and use the bounding box to crop the image. The cropped image is fed into the segmentation process. In the detection process, we use MaskRcnn [3].

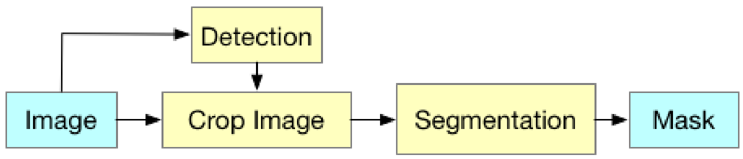

Figure 1: Process

In the segmentation process, we design an encode-decode architecture of network inspired by DeepLab [4] and PSPNet[5] for segmentation. Our architecture is shown in Figure 2. The features are extracted by ResNet 101. We also cascade three blocks after ResNet 101 with setting dilation of convolution to two. After ResNet, the feature maps are feed into a multi-scale block. In this block, we use different size of pooling and dilation convolution to extract different scale features. Those feature maps are concatenated with each other and upsample to input size.

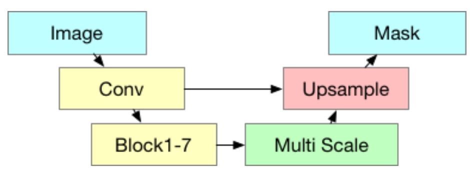

Figure 2: Architecture

Figure 3 shows the architecture of the multi-scale block which combines pooling layers and convolution layers. We use three pooling whose size is 3,5,7 separately and three dilation convolutions whose dilation is 3,6,12 separately. The outputs of those layers are concatenated to feed into the upsample block.

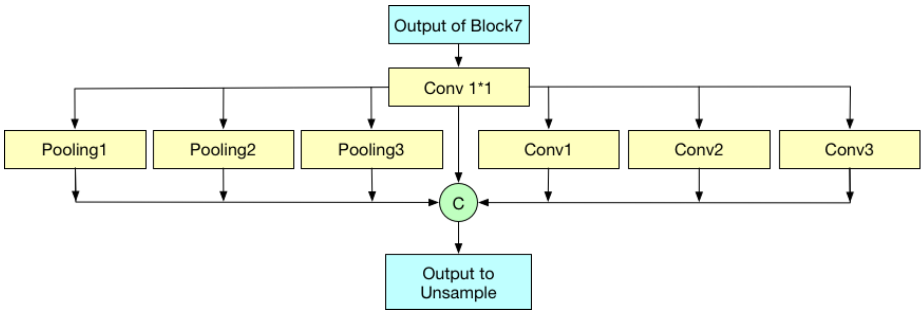

Figure 3: Multiscale

Figure 4 shows the architecture of the upsample block. The output of the multi-scale block is unsampled by 4 and concatenate with low feature map from ResNet. After that, the feature map is unsampled to the input image size.

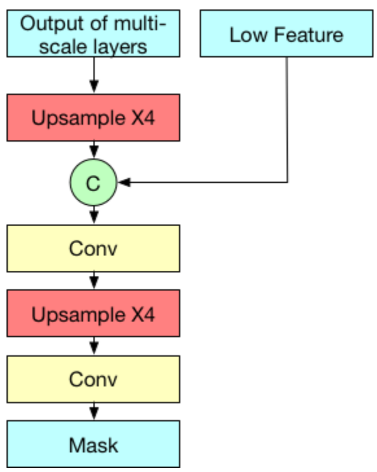

Figure 4: Upsample

2.3 Pre-processing

We do some pre-processing to improve the performance of our method. Instead of only using RBG channels, we combine the SV channels in Hue-Saturation-Value colour space and lab channels in CIELAB space with the RGB channels. We also do some rotation, colour jitter, flip for the input images. Each channel of images is scaled in 0 to 1 and the size of images is resized to 512x512. Some examples are shown in Figure 5.

(a) Flip Left to Right (b) Roatate 45°

Figure 5: Examples for Image Augmentation

2.4 Training

We use the Adam optimisation and set the learning rate to 0.001. The learning rate will be set at 92% after each epoch. The batch size is 8. The total epoch is set to 400. We early stop the training when the net starting overfitting. We use the dice loss which is shown in Equation 1, where $p_{i,j}$ is the prediction in pixel (i, j) and $g_{i,j}$ is the ground truth in pixel (i, j). Compared with cross-entropy loss, dice loss could consider more global information about the lesion.

$$L = -\frac{\sum_{i,j}(p_{i,j}g_{i,j})}{\sum_{i,j} p_{i,j} + \sum_{i,j} g_{i,j} - \sum_{i,j}(p_{i,j}g_{i,j})} \tag{1}$$

2.5 Implementation

Our method is implemented by Pytorch 0.3.1 in Ubuntu 14.04. We use two Nvidia 1080ti with 11 GB.

3 Results

In ISIC2018 validation set, our IOU is above 0.78.

We thanks for the supporting and advice from our teammates in MTLab.

References

[1] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci. Data*, 5:180161, 2018.

[2] Noel C. F. Codella, David Gutman, M. Emre Celebi, Brian Helba, Michael A. Marchetti, Stephen W. Dusza, Aadi Kalloo, Konstantinos Liopyris, Nabin K. Mishra, Harald Kittler, and Allan Halpern. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on

biomedical imaging (isbi), hosted by the international skin imaging collaboration (ISIC). *CoRR*, abs/1710.05006, 2017.

[3] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross B. Girshick. Mask R-CNN. *CoRR*, abs/1703.06870, 2017.

[4] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L. Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *CoRR*, abs/1606.00915, 2016.

[5] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. *CoRR*, abs/1612.01105, 2016.