Summer Internship Report
Russel Shawn Dsouza

Contents

Semantic Segmentation of H&E stained nuclei in Histopathology images

Abstract – Medical imaging results in faster assessment of almost every disease from lung cancer to heart disease. By automating the analysis of medical images we can unlock even faster diagnosis of critical diseases from cancer to common-cold. The detection and segmentation of cell nuclei in slide images is a major part of most automated detection & diagnosis methods.

1 Introduction

Tissue slides are informative on many aspects of the disease such as cancer subtype, the grade or the reaction of the patient's immune system, morphological properties of cells and tissues, spatial organization of cells to name a few. However, interpretation of such data is not straightforward and requires special training and a deep understanding of the underlying physiology.

The revolutionary step of tissue glass slide digitization has opened up exciting possibilities in the world of digital pathology. With recent advancements in digital pathology, a high volume of quality digitized data is available for algorithm developers, scientists, and pathologists around the world.[1]

2 Literature Review

Sliding Window Approach for detection and segmentation of nuclei in H&E images CNNs FCN[2] Postprocessing

3   Problem Statement

Segment the cell nuclei from the surrounding tissue in H&E stained Histopathology images of kidney tissues using the U-Net architecture.[3]

4   Dataset

H&E stained images of kidney tissues was used as the dataset for this project.

The dataset is pre-divided into 3 sections: Train, Validate, Test. The Train set has 400 $400 \times 400 \times 3$ images, where 3 is the number of channels and refers to the RGB colorspace. Each image has a unique ID and the corresponding ground-truth to each image is also names using the same ID. The ground-truth images are of the size 400×400 i.e. grayscale images with only 0 or 255 intensities. The Validate set has 50 images with ground-truth, the same size as those in the Train set. The Test set has 20 images, the same size as those in the Train set.

5   Architecture

U-Net[3] Color deconvolution[4] Color normalization

6   Implementation

- PyTorch[5]

- Kaiming Initialization

- Data Augmentation

- torchvision

- morphological postprocessing with opencv

- metrics in sklearn/ignite

7   Experimental results and Discussions

8   Conclusion

9   Future Scope

References

[1] G. Kamble, "Deep learning in digital pathology."

[2] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *CVPR*, 2015.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *ArXiv*, vol. abs/1505.04597, 2015.

[4] A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution.," *Analytical and quantitative cytology and histology*, vol. 23 4, pp. 291–9, 2001.

[5] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," in *NIPS Autodiff Workshop*, 2017.