CELL AND SUB-CELLULAR SEGMENTATION IN QUANTITATIVE PHASE IMAGING USING U-NET

Jakub Majerčík, Michal Špaček
Bachelor Degree Programme (3), FEEC BUT
E-mail: xmajer19@stud.feec.vutbr.cz, xspace31@stud.feec.vutbr.cz

Supervised by: Tomáš Vičar, Jaromír Gumulec
E-mail: vicar@vutbr.cz, j.gumulec@med.muni.cz

Abstract: The ability to automatically segment images, especially microscopy images of cells, opens new opportunities in cancer research or other practical applications. Recent advancements in deep learning enabled for effective single-cell segmentation, however, automatic segmentation of sub-cellular regions is still challenging. This work describes an implementation of a U-net neural network for label-free segmentation of sub-cellular regions on images of adherent prostate cancer cells, specifically PC-3 and 22Rv1. Using the best performing approach, out of all that have been tested, we have managed to distinguish between objects and background with average dice coefficients of 0.83, 0.78 and 0.63 for whole cells, nuclei and nucleoli respectively.

Keywords: cell segmentation, deep learning, neural network, quantitative phase imaging, nucleus, nucleolus

1 INTRODUCTION

Automatic separation of objects from the background is often an inseparable part of computer vision-related problems and finds its use in many medical and clinical applications. Accurate segmentation of cells can speed up the whole research process of many diseases, such as cancer. By identifying different cellular compartments, we are able to analyze the biological processes taking place in these cells. In this work, we focus on the label-free segmentation of prostate cancer cells in quantitative phase refractive index microtomograms (PC-3 and 22Rv1) using machine learning methods. Our goal is to segment whole cells, nuclei, and nucleoli with reasonable accuracy. Correct segmentation of sub-cellular regions is a necessary step for automated sub-cellular cell analysis.

2 METHODS

2.1 CELL CULTIVATION

In this work, we have analysed 2 cell lines. The PC-3 human epithelial cell line was obtained from fourth-degree prostate adenocarcinoma and bone metastases. 22Rv1 is a human prostate cancer epithelial cell line obtained from a xenograft of propagated mice after castration. Both PC-3 and 22Rv1 cells were cultured at the Institute of Pathological Physiology, Faculty of Medicine, Masaryk University. PC-3 cells were cultured in Ham's F12 medium with 10% FBS, while RPMI-1640 medium with 10% FBS was used to culture 22Rv1 cells. Both media were enriched with antibiotics (penicillin 100μ/ml streptomycin 0.1mg/ml). The cells were maintained in an incubator at a constant temperature of 37°C, 60% humidity with 5% CO_2 (Sanyo, Japan). Passages of 22Rv1 cells ranged from 25 to 35, the passage of the PC-3 line from 15 to 25. Cells showing 50-60% confluence were washed with an FBS-free medium. Subsequently, a fresh medium containing FBS was added. Furthermore, variants of 22Rv1 and PC-3 cells adapted to high zinc concentrations were used in this study to increase data variability, previously prepared by Holubova et al [2].

2.2 DATA ACQUISITION AND ANNOTATION

The acquisition of microscopic images was performed using a Nanolive 3D Cell Explorer holotomographic microscope with a 60x/0.8 objective, which measures the quantitative refractive index. Thus, it enables the noninvasive acquisition of 3D data of almost any cell type [1]. The refractive index of both media was measured at 1.34. All four cell types (PC-3 Wild type, PC-3 Zinc resistant, 22Rv1 Wild type and 22Rv1 Zinc resistant) were cultured in separate ibidi μ-Slide Luer 0.8 chambers. Thanks to this, it was possible to easily create 4 unique data sets containing 200-220 3D images. Each image is composed of 96 slices and, upon acquisition, was identified by a label number in which the individual cell structures could be seen the best. In addition to the microscope, the acquisition also required software, specifically *Steve*, provided by Nanolive. The dimensions of the field of view were 94x94x35μm. The measurement resulted in more than 750 usable 3D images ($\sim$400 PC-3 and $\sim$350 22Rv1) containing more than 1500 cells.

The next step in creating our dataset was annotation of the measured images. Every image was manually segmented, which yielded segmentation masks of the whole cell, nucleus and nucleolus. These masks will later serve as so-called *ground truth* images in the process of training the neural network. Since the image resolution in 3rd dimension is relatively low and using all 96 slices as an input would be computationally costly, it was suitable to convert the original 512x512x96 3D image to a single-channel grey-scale 2D image: (1) by averaging through the 3rd dimension, (2) by maximum intensity projection over the 3rd dimension, (3) by standard deviation projection over the 3rd dimension and (4) by selecting the slice with the highest entropy value, thus the sharpest one. The data modified in this way contains most of the information needed for segmentation purposes.

2.3 DATA PREPROCESSING

The data had been normalized using Z-score to ensure its homogeneity. Every type of a converted 2D image was normalized separately with the corresponding values of mean and standard deviation computed through the whole dataset. Judging by the fact that segmentation of sub-cellular regions is a relatively complex task and our dataset is not large enough, the implementation of augmentation is essential to acquire the ability to train extensively and thus achieve the desired results. Our augmentation algorithm consists of: (1) addition of a random number from a predefined range, (2) multiplication by a random number from a predefined range, (3) affine transformation (random rotation, translation, shear and tilt), (4) random blurring and sharpening. All of these augmentation parameters had been set based on prior testing. After augmentation, the input image was randomly cropped to the size of 256x256 pixels.

2.4 IMPLEMENTATION

All mentioned algorithms were implemented in Python using a deep learning library PyTorch. For the purposes of this work, the neural network of choice was U-net [3], the most widely used architecture for biological image segmentation, as well as a U-net pretrained on an ImageNet database with a ResNet18 encoder [4] to compare to and possibly improve the results. Both networks have been trained on all 4 of the above-mentioned 2D images separately in addition to a stack of these 4 images. To properly segment all 3 regions (cell, nucleus and nucleolus), it was necessary to train 3 separate models, 1 for each region and input image type. This resulted in 30 trained models in total. Every model was trained with the same parameters set based on thorough testing and optimization.

The data-set was split in a 4:1 ratio; thus, 80% of the images were used for training and 20% for testing. The number of training iterations was set to 1000 with a batch size of 16. The network's parameters were optimized by the Adam optimizer. The scheduler had been reducing the initial learning rate of 0.0005 to 10% of its prior value at iterations number 500, 750 and 900. The model's perfor-

mance during training was measured by dice loss, computed against manually segmented ground truth images. The whole training algorithm ran on the remote server in the Department of Biomedical Engineering, Brno University of Technology with following computer specifications: Intel(R) Core i7 3.07 GHz CPU with an 8GB GDDR5 NVIDIA GeForce GTX 1070 GPU.

3 RESULTS

The performance of all models have been evaluated after training on testing data with no cropping applied (512x512 input image size) and by the same metric as during training - dice coefficient. Average values are shown in Table 1, from which the following conclusions can be made: (1) different 2D input image types seem to have only a negligible effect on the final performance, (2) stack of all 4 2D images as an input does not improve the performance, and (3) pretrained network allowed for a substantial improvement in nuclei and nucleoli segmentation and slight improvement when segmenting whole cells. Based on these results, the most useful approach is to use the slice with the highest entropy value as an input for the pretrained network. Utilizing mentioned approach, dice coefficients for both cell lines have been computed separately, values in Table 2. It is clear that segmentation of the 22Rv1 cell line is more problematic, with notable drop in nucleoli segmentation performance. Figure 1 shows examples of PC-3 and 22Rv1 cells with binary segmentation masks compared to ground truth masks.

Although the coefficients do not suggest an outstanding performance, we believe that the network learned to segment cellular structures reasonably well. Many nuclei and nucleoli cover only a small number of pixels. This leads to a relatively high error when the manually segmented mask does not cover these regions perfectly and the network predicts it correctly. Secondly, considering that no staining had been applied when measuring the cells (label-free), the results are reasonably high.

Table 1: Average dice coefficients for various types of input images and sub-cellular regions on testing data using both networks.

	U-net trained from scratch			**Pretrained U-net**		
Input image type	Whole cell	Nucleus	Nucleolus	Whole cell	Nucleus	Nucleolus
Mean	0.75	0.59	0.44	0.84	0.79	0.60
Max	0.80	0.61	0.50	0.83	0.77	0.59
Entropy	0.79	0.62	0.53	0.83	0.78	0.63
Std	0.80	0.56	0.49	0.84	0.78	0.61
Stack of all 4	0.79	0.57	0.51	0.83	0.78	0.62

Table 2: Average dice coefficients for PC-3 and 22Rv1 cell lines and their sub-cellular regions on testing data using pretrained network with the highest entropy slice as an input.

Cell line / Region	Whole cell	Nucleus	Nucleolus
PC-3	0.86	0.83	0.73
22Rv1	0.79	0.71	0.51

4 CONCLUSION

Our work was focused on the automatic segmentation of 2 prostate cancer cell lines using deep learning methods. The proposed algorithm manages to distinguish between cells, nuclei, nucleoli and background reasonably well, considering the fact no specific staining was used (label-free imaging). Our effort to achieve better results revealed that the use of a pretrained model leads to a substantial

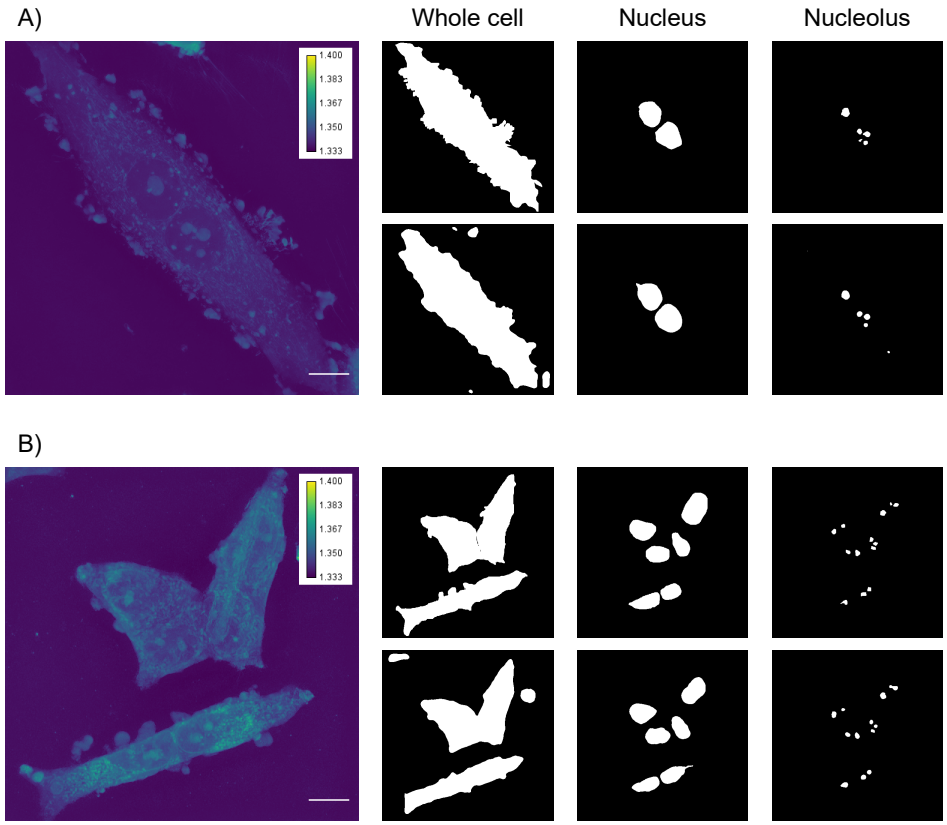

Figure 1: Examples of A) PC-3 and B) 22Rv1 cells, quantitative phase imaging, max projection, scale bar indicates 10 μm; their corresponding binary ground truth masks (top rows) compared to binary masks generated by the best performing model (bottom rows), i.e. pretrained U-net with a slice with the highest entropy value as an input.

performance improvement, especially for nuclei and nucleoli, 21% and 16% respectively. The process of segmentation is often necessary for subsequent analysis, and our trained models eliminate the need to segment these images manually. This allows for automation and potentially removing human error. Our subsequent work will utilize this algorithm and generate masks to quantitatively analyze the cellular properties and characteristics of the mentioned cell lines.

REFERENCES

[1] Live Cell Imaging Microscope: 3D Cell Explorer. Dec 2020.
URL `https://www.nanolive.ch/products/3d-microscopes/cx/`

[2] Holubova, M.; Axmanova, M.; Gumulec, J.; aj.: KRAS NF-κB is involved in the development of zinc resistance and reduced curability in prostate cancer. *Metallomics*, ročník 6, č. 7, 2014: s. 1240–1253.

[3] Ronneberger, O.; Fischer, P.; Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, s. 234–241.

[4] Yakubovskiy, P.: Segmentation Models Pytorch. 2020.
URL `https://github.com/qubvel/segmentation_models.pytorch`