Multi-organ Nucleus Segmentation Using UNet++

Xuhui Guo, Xinru Song, Yixuan Deng, Zhiran Wang

1 Introduction and Problem statement

The major component of this project is basically a semantic segmentation task for the nucleus in H&E stained histology images of multi-organ tumor cells. Accurate segmentation of cell nuclei allows the analysis of key features like density, size ratios, and shape variations. These are critical for evaluating cancer severity and predicting treatment outcomes. In this project, we will employ UNet++[5] model to perform this task. UNet++ is a new general purpose image segmentation architecture for more accurate image segmentation. This architecture integrates a series of UNets [4] with variable depths. These redesigned skip pathways are strategically devised to surmount two predominant limitations associated with the conventional UNet architecture: firstly, the indeterminate optimal depth, and secondly, the overly constricted configuration of skip connections.

2 Method

The major difference between UNet and UNet++ is the design of their skip connection. Due to the dissimilarities between the same-scale feature maps from the encoder and decoder in the UNet skip connections, the best match for feature fusion cannot be guaranteed. Hence, UNet++ is motivated to reconstruct its skip connection by letting itself consist of an encoder and decoder through a series of nested dense convolutional blocks. There are another two methods adopted to design UNet++. First, deep supervision was introduced. Then, model pruning was enabled as a result of deep supervision. Therefore, UNet++ embeds UNet of different depth in its architecture, which addresses the unknown optimal depth limitation. Moreover, UNet++ gets rid of the restrictive skip connections, meaning that only the same-scaled feature maps can be fused. Hence the second limitation gets addressed.

3 Related work

Image segmentation has been a popular area worthing exploring. Fully convolutional network[3] was first proposed in 2015, based on the fundamental idea on image classification with CNN, they developed a "fully convolutional" network which trained end to end, pixels to pixels on semantic segmentation for pixelwise prediction. UNet[4] was also another widely used architecture in computer vision specifically for image segmentation tasks. By employing a u-shaped architecture with an encoder and decoder and utilizing skip connection between the corresponding layer both in the encoder and decoder,the architecture enables fine-grained spatial information. The other recent related works are GridNet[1] and Mask-RCNN[2]. GridNet is an encoder-decoder where the feature maps are connected on a grid fashion while Mask-RCNN is critical framework for object detection, classification and segmentation where UNet++ can be deployed as the backbone architecture in.

4 Novel Component

We will implement UNet++ on the MoNuSeg 2018 challenge dataset which was obtained by carefully annotated tissue images at the cellular level of tumors from different organs. The datasets provides pixel-level annotations for slices of various tissue structure and components, making it valuable to explore fine-grained image segmentation on medical images. This dataset typically contains high-resolution histopathological images with intricate tissue structure while UNet++, a well designed architecture for handling high-resolution images due to its dense concolution in skip connection, is suited for this task.

5 Plan of Experiments

5.1 Implement

We will generally follow the official PyTorch implementation repository built by the author of the UNet++ paper[5]. Since the repository was adapted from the newly published repository of nnUNet which covers parts that are required for a complete library but will not be utilized by us. Therefore, one major task is to recover the plain code for model structure and training section.

5.2 Datasets

The dataset is provided by the MoNuSeg 2018 challenge dataset which can be downloaded from `https://monuseg.grand-challenge.org/Data/`. The training set consists of 30 H&E stained tissue images captured at 40x magnification from TCGA archive and around 22,000 nuclear boundary annotations. Due to the dataset being insufficiently large, we will use a sliding window mechanism to extract 96×96 patches from the images, and randomly assign them into training set (70%), validation set (15%), and test set (15%).

5.3 Evaluation

In the nuclei segmentation task, the foreground (nucleus) is sparse compared to the background. Therefore, we will use mean Intersection over Union (mIoU) and the Dice coefficient as evaluation metrics. mIoU measures the ratio of the intersection and the union of the predicted and ground truth segmentation maps, averaging across all classes, which mitigates the influence of imbalance. While the Dice coefficient measures the overlap between the predicted and true segmentation maps, defined as twice the intersection area divided by the sum of the predicted and ground truth segmentation areas. It is particularly suitable for medical imaging tasks like nucleus segmentation, capturing spatial overlap effectively.

6 Plan of Project

Our target of this project is producing the code based on the paper and applying it on our own novel dataset.

6.1 Anticipated division of work

We will primarily divide the content of this project into four parts. 1. Data Preprocessing: Augment the dataset and assign them into training, validation, and test sets. 2. Model Construction: Implementing the UNet++ architecture based on the description and code from the paper. 3. Training and Evaluation: Training our model using the dataset and evaluating its performance, followed by further adjustments. 4. Final Report: Documenting the entire project process and showcasing the final results. The specific allocation of tasks among team members will be adjusted based on individual preferences and interests.

6.2 Experience and Expertise

This project requires us to utilize expertise from multiple domains. Firstly, we need to apply knowledge of data preprocessing to handle images. Next, we need to further study the UNet++ model and appropriately utilize tools like PyTorch to construct neural networks.

6.3 Challenges and Contingency Plans

The main challenges we face are the limited training data and the difficulty in architecting the UNet++ model without detailed guidance.

If we encounter difficulties in architecting the model, we will consider seeking help from relevant communities or forums associated with UNet++ and cell segmentation.

6.4 Milestones

Week 11: Finishing data preprosessing and start working on model construction.
Week 12 - Week 13: Model Construction.
Week 14 - Week 15: Finishing model construction and start training, hyperpaprmater tuning, and evaluating the model.
Week 16: Finishing final report.

References

[1] Damien Fourure, Rémi Emonet, Élisa Fromont, Damien Muselet, Alain Trémeau, and Christian Wolf. Residual conv-deconv grid network for semantic segmentation. *CoRR*, abs/1707.07958, 2017.

[2] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[3] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.

[5] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE transactions on medical imaging*, 39(6):1856–1867, 2019.