Glomerulus Segmentation in Kidney Biopsies

Computer Vision Project, Group 9

Lukas Göbl, Peer Schäfer, Lukas Scheib

15.07.2025

Problem Statement

General: Accurate segmentation of glomeruli in kidney biopsy images is crucial for diagnosing chronic kidney diseases.

Precise Tasks: Segment glomeruli at patch-level and whole-slide level in kidney tissue images as part of the MICCAI 2024 KPIs Challenge.

WSI Example: train\56Nx\12_116

Images/wsi_thumbnail_with_grid.png

WSI Mask Example: train\56Nx\12_116

Images/mask_thumbnail_with_grid.png

VariableUNet Model

- Standard UNet architecture with double convolution layers and skip connections.
- Dilated bottleneck for multi-scale feature extraction.[1]
- Channel counts scale with variable patch size.
- Combined BCE and Dice loss (BCEDiceLoss) to address class imbalance. Focusing on recall and F_1-score.

[1] Li et al. (2021)

Experimental Design

- Trained on official KPIs dataset (train/val/test splits).
- Hyperparameter tuning: BCE/Dice ratio, dice smooth, positive class weighting for BCE, base learning rate of Adam optimizer.
- Experiments with different patch sizes (256, 2048 px) and data subsets.
- Evaluated using accuracy, precision, recall, and F_1-score.
- Qualitative assessment via visual inspection.

Metrics of our best performing models

Patch size	Split	Accuracy	Precision	Recall	F_1-score
256	Validation	0.9781	0.7479	0.7660	0.7569
	Test	0.9757	0.7537	0.7133	0.7330
2048	Validation	0.9705	0.9299	0.3662	0.5255
	Test	0.9643	0.9159	0.2616	0.4070

Hyperparameters:

- BCEDiceLoss Ratio $= 0.8$
- BCE Positive Weight $= 25$
- Dice Smooth $= 10^{-6}$
- Learning Rate $= 10^{-4}$

WSI Segmentation

Implemented patching pipeline with the following hyperparameters:

- Patch size
- Patch overlap (stride)
- Trivial patch threshold for positive pixels

Possible next steps:

- Implementation of a stiching pipeline to combine patch predictions into a full WSI mask.
- Train VariableUNet with complete patching and stiching pipeline.

Discussion

What worked:

- BCE + Dice loss crucial for recall and F_1-score.
- Hyperparameter tuning improved glom detection.
- WSIs can be patched for patch-level segmentation.

Limitations:

- Instability in training, especially with full-size patches.
- Our configuarations, trained on small patches, did not generalize well to larger patches.
- No empirical validation on WSI segmentation due to time constraints.

Discussion

In Retrospect:

- Greater emphasis on pre- and post-processing (downsampling patches and upsampling predictions).
- More in-depth data exploration regarding samples of different deseases.

Future Work:

- Exploring different model architectures and methods.
- Search for more robust hyperparameters and domain adaptation.
- Exploring different WSI patching thresholds.

Appendix

UNetArchitecture

Images/UNetArchitecture.png

Figure: Doi: 10.1117/1.JMI.8.6.067501

Double Convolution

```python
class DoubleConv(nn.Module):
    def __init__(self, in_channels, out_channels):
        super().__init__()
        self.double_conv = nn.Sequential(
            nn.Conv2d(in_channels, out_channels, 3, padding=1),
            nn.BatchNorm2d(out_channels),
            nn.ReLU(inplace=True),
            nn.Conv2d(out_channels, out_channels, 3, padding=1),
            nn.BatchNorm2d(out_channels),
            nn.ReLU(inplace=True),
        )

    def forward(self, x):
        return self.double_conv(x)
```

Heat Map F1 Score

Images/12_heatmap_f1_score_by_weights.png

Heat Map Recall

Images/12_heatmap_recall_by_weights.png

FlipFlop Graph

Images/Exp3_acc_prec_rec.png