Glomeruli Functional Tissue Units Detector in Kidneys Using U-Net and ResNet

Abstract

The domain of healthcare is a crucial need of today's era. Of late, modern-day technologies such as machine learning, deep learning, and artificial intelligence are being vastly employed to improve the health condition of humans in society. Curing millions of diseases, analyzing infections, providing accurate test results, and performing high-level maintenance checks are now all possible to be carried out with ease with the evolution of technology. The research sector has been working at a continuous pace to accomplish various studies that identify different organs and assess their working mechanisms in the human body. One such study includes the identification of glomeruli functional units in the human kidney tissue. Each glomerulus is a small ball-like structure that is composed of blood vessels and performs the filtration of the blood to form urine. The paper presents two deep learning-based models for detecting them in kidney tissue images. The study will help in understanding the relationships between cell and tissue organization and functioning. It can be used by researchers in cell and tissue anatomy to develop diagnostic techniques. After implementation, the proposed model obtained an accuracy of 94.73% with a Dice Coefficient of 0.825 for U-Net and accuracy of 95% with a Dice Coefficient of 0.843 for ResNet50.

1. Introduction

Good health of the human body is one of the crucial needs of this society. It is required to analyze the human body, study different functionalities of the human organs to discover the diseases and treatments needed for curing them once affected. Thus, technology came into play in the health sector that has been helping in revolutionizing it. Machine learning, deep learning, and artificial intelligence have been applied to this sector for analyzing and gaining knowledge, diagnosing diseases, or developing methods to cure these diseases.

The kidneys are the most important organ in the human body for the proper functionality of the excretory system. The main function of kidneys is to filter and remove all the waste products and excessive fluids from the blood in the form of urine. The functional unit of the kidney is known as the nephron, which consists of glomeruli. These glomeruli help the kidney to filter out the urine, which is the excessive or waste fluid present in the human body, from blood. These functional tissue units (FTU) are the blocks of cells centered around a capillary, such that every cell of this block is within a certain distance from any other cell in the same block. Damage to glomeruli can cause a variety of diseases such as albuminuria- large amounts of protein in the urine, hematuria- blood in urine, hypoproteinemia- low blood protein, edema- swelling in certain parts of the body, and even renal failure in the long term[1]. Further, the glomerulus identification and classification into sclerotic/non-sclerotic is important for every kidney transplant and can help in analyzing the structural patterns and health of the kidney. When done traditionally under a microscope by pathologists, it can be an extremely tedious and error-prone procedure. Hence, the application of machine learning techniques in this field is an evolving research area that could add significant value to nephrological sciences. [2]

This paper presents deep learning-based models for the identification of glomerular FTUs from images of different human kidneys. We used the PAS-stained kidney image dataset obtained from the Human BioMolecular Atlas Program (HuBMAP) dataset. This dataset was developed by researchers and biologists of the HuBMAP organization for the development of a framework to map the human body at a single-cell resolution. The dataset consists of fresh frozen and formalin-fixed paraffin-embedded (FFPE) Periodic acid-Schiff(PAS) stained kidney images in the TIFF format. Thus, after applying two deep learning-based models for identifying the glomeruli structures present in the human kidney from images, the best model achieved was Resnet with the accuracy of 95% and a Dice Coefficient of 0.843, proving that deep learning models are an efficacious tool for identifying glomeruli in kidney tissue images.

2. Related Work

In the past, significant work was done for the identification of glomerular FTUs in the human kidney. In 2021, Bouteldja et al. proposed a deep learning algorithm to provide accurate segmentation of PAS-stained kidney tissues obtained from healthy mice, five murine disease models,

and from other species used in preclinical research. They investigated the use of convolutional neural network architecture for this task. They trained the CNN to segment six major structures: glomerular tuft, glomerulus including Bowman's capsule, tubules, arteries, arterial lumina, and veins [3]. The trained CNN yielded high performance in most species used in research—including rats, pigs, bears, and marmosets—as well as in humans.

In 2021, Jayapandian et al. [4] proposed a study on the development and application of deep learning-based automatic segmentation of the histologic structures in kidney biopsies and nephrectomies. The study includes the classification of renal histologic primitives on four different stains using 125 biopsies from the Nephrotic Syndrome Study Network (NEPTUNE) dataset. After applying deep learning techniques, it was concluded the best results were obtained from PAS-stained whole slide images while silver-stained whole slide images yielded the worst performance. Owing to the outcome of this study, all studies that followed have preferred using PAS-stained kidney images for building their models.

A study carried out by Bueno et al. [5] focused on the identification and characterization of glomeruli in kidney diseases using the comparison of U-Net and SegNet CNN models to identify the glomerulosclerosis in the human body. The algorithms were implemented over a dataset composed of 47 whole slide images of human kidneys stained with PAS. SegNet for two-class segmentation followed by a fine-tuned AlexNet network generated the best results with an accuracy of 98.16%.

Zeng et al. [6] performed their study on the identification of glomerular lesions and intrinsic glomerular cell types in diseased kidneys. For the identification, the study proposed a computer-aided system using deep learning techniques that obtained an accuracy of 92.12%. The results indicated the ratio of mesangial cells, endothelial cells, and podocytes within glomeruli to be 0.41: 0.36: 0.23 with a good performance score.

Uchino et al. [7] worked on developing a method for the classification of glomeruli using a nephrologist-AI collective intelligence approach on a dataset of 283 kidneys biopsies, consisting of 15,888 glomerular images from various nephrologists. After analyzing, the model obtained an area under the curve score of 0.986 for PAS and 0.983 for periodic acid methenamine silver. The study concluded that advanced techniques can be used to improve the diagnostic accuracy of the system.

3. Approach

3.1. Image Segmentation

With the evolution of technology, it has now become possible to detect and determine the boundary or shape of an object or its orientation from its images. This is called image segmentation. Formally, it is the technique where a pixel-wise mask is formed around the object that needs to be classified in the input image[8]. It can be performed by considering every pixel of the image. Wherever pixels of the same intensity are found, the model considers that cluster of pixels to be a part of the result list and produces a boundary over the object contained by them to identify it.

3.2. Glomerulus Detection Problem

Our approach models the problem of glomerulus detection using image segmentation as a pixel-level classification problem where we classify each pixel to one of the two classes- 'glomerulus'and 'not glomerulus', respectively denoted by 1 and 0 in the result.

3.3. Deep Learning Models

Deep learning is a sub-field of machine learning where the algorithms mimic the structure and function of the human brain using artificial neural networks. With the advancement in technology, deep learning has been having a major impact on the healthcare industry. It has been successfully generating granular summary, analyzing the human body through images and test reports, thus solving a variety of disease detection and curing problems. By using the image segmentation technique along with deep learning techniques, it is possible to detect the glomeruli present in human kidneys from a variety of kidney images, that can be further helpful in identifying different diseases related to human kidney. Traditional convolutional neural networks have been successful for simpler images all these years but haven't given good results for complex images. This is where the advanced architectures like U-Net and ResNet come into play.

U-Net: U-Net is a convolutional layer network which was especially developed for solving the problem of image segmentation in biomedical domain[9]. The network is based on the concept of fully convolutional networks introduced by J. Long et. al[10]. U-Net can be trained end-to-end using very few images and outperforms the sliding-window CNNs. It is highly efficient in executing pixel-wise classification and regression learning. It consists of Convolution Operation, Max Pooling, ReLU Activation, Concatenation, and Up Sampling Layers and three sections: contraction, bottleneck, and expansion section. U-Net uses a loss function for each pixel of the image. This helps in easy identification of individual cells within the segmentation map. Softmax is applied to each pixel followed by a loss function.

ResNet: Residual Network(ResNet) is a popularly used neural network for many challenging computer vision tasks. Traditional deep networks are hard to train because of the vanishing gradient problem. The ResNet architecture is in-

spired by the pyramidal cells in the cerebral cortex of the human brain. In traditional neural networks, each layer feeds into the next layer. But in a network with residual blocks, each layer feeds into the next layer and directly into the layers about some hops away. In other words, ResNet gets rid of the vanishing gradient problem by utilizing skip connections, or shortcuts to jump over some layers so that the gradient during backpropagation is maintained. ResNet50 is a ResNet model with 48 convolution layers along with 1 MaxPool and 1 Average Pool layer.

3.4. Optimizers

Adam: We have used the **Adam optimizer** as an adaptive learning rate optimizer for our U-Net experiments. Adam optimizer was introduced by Diederik Kingma et al.[11]. Adam is essentially a replacement optimization algorithm to SGD for training deep learning models. Adam combines the best properties of the AdaGrad and RMSProp algorithms to provide an optimization algorithm that can handle sparse gradients on noisy problems[12]. Also, majority of experiments on biomedical image segmentation and classification including those by Muhammad Yaqub et al.[13] have concluded that Adam optimizer is the best suited optimizer for this family of tasks.

AdamW: We have used both Adam and AdamW optimizers for our ResNet experiments. AdamW is an improvement over the Adam optimizer that yields better training loss and more generalized models compared to models trained with Adam. AdamW decouples the weight decay from the optimization step, allowing the model to optimize the weight decay and the learning rate separately[14].

3.5. Metrics

3.5.1 Dice Loss

One of the metrics we are using for evaluating our models is the Dice Coefficient. The Dice Coefficient is being calculated for every image in the dataset, by comparing the original segmentation of the pixels(provided via the mask) and the segmentation result returned by our model for that image. Dice Loss and Dice Coefficient originate from Sørensen–Dice Coefficient, which is a statistic developed in 1940s to gauge the similarity between two samples. The Dice Coefficient can be used to compare the pixel-wise agreement between the predicted segmentation and its corresponding ground truth. The mathematical formula of the same is:

$$\frac{2*|X \cup Y|}{|X|+|Y|}$$

where X is the set of pixels predicted as glomeruli and Y is the actual set of pixels containing glomeruli. The Dice Coefficient is defined to be 1 when both X and Y are empty. Dice Loss can be computed as 1 - Dice Coefficient.

3.5.2 Accuracy

The pixel-wise accuracy metric in our experiments is being calculated for every image using the following formula:

$$\frac{tp+tn}{tp+tn+fp+fn}$$

where for a particular image, tp denotes the number of pixels correctly classified as 'glomerulus', tn denotes the number of pixels correctly classified as 'not glomerulus', fp denotes the number of pixels wrongly classified as 'glomerulus', and fn denotes the number of pixels wrongly classified as 'not glomerulus'.

3.5.3 Relation Between Metrics

We track the Dice Loss and the Accuracy metrics for our models, to evaluate their performance. While both these metrics are correlated, the accuracy metric is more intuitive for gauging the proportion of the complete area of an image that was correctly segmented by the model including both 'glomerulus'and 'not glomerulus'pixels. On the other hand, the Dice Loss compares the sets of pixels classified into 'glomerulus'in the output and the target respectively.

3.5.4 Preprocessing and Batch Size

The TIFF images provided in the dataset each spanned number of pixels in the order of 10,000 both horizontally and vertically. For making the training more convenient, we sliced every image to multiple images of dimension 512 $\times$ 512 each. Appropriate padding was applied to the images and respective targets in order to achieve uniform sizes. The batch size was set to 8 for the U-Net implementations and 32 for the ResNet50 implementations.

4. Experiment

The HuBMAP Kidney Segmentation dataset was used for training, predicting pixel-wise classes, and computing performance metrics on the identification of glomerulus in human kidney. The data was thoroughly analyzed before being used for training and testing on the models. Before training the models, the input images were processed by adding some padding and were resized to get consistent width and height. The training dataset consisting of RLE-encoded annotations was used to generate binary mask where 1 indicates that pixel contains 'glomerulus 'while 0 indicates that pixel does not contain 'glomerulus '. Then these masks and images were used for training the U-Net implementation provided by the Keras library, and the ResNet50 implementation provided by the PyTorch library. K-fold cross-validation is used for preventing over-fitting on a single validation set. The training and evaluation of all the models were performed independent of each-other.

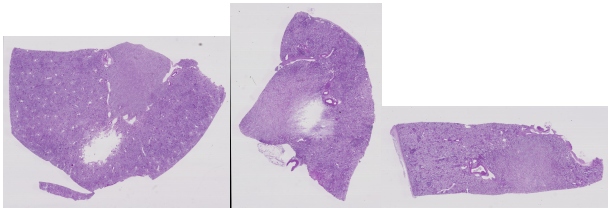

Figure 1: Sample images from datasets

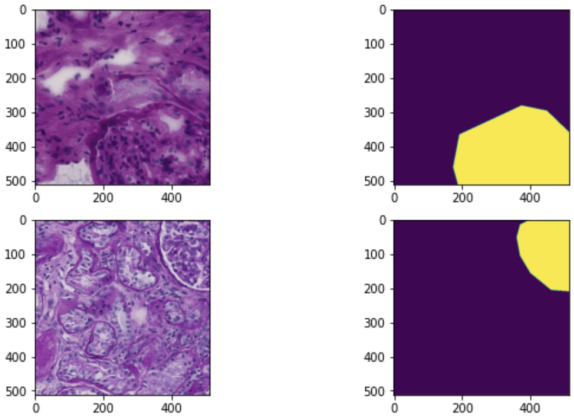

Figure 2: Preprocessed training image sample on left and mask on right

4.1. Dataset

The HuBMAP Kidney Segmentation dataset includes a mix of fresh-frozen and formalin-fixed paraffin-embedded(FFPE) PAS-stained kidney images [15]. The HuBMAP dataset is comprised of very large (≥500MB - 5GB) TIFF files which are divided into two subsets: training and the test datasets which are used sequentially for the training of model and the identification of glomeruli.

The HuBMAP training dataset consists of 15 training images, along with corresponding annotations in both RLE-encoded and unencoded(JSON) formats where these annotations represent classification of each pixel into 'glomerulus'and 'not glomerulus'. These files are in the TIFF format so as to provide better resolution and clarity of the images. Figure 1 contains a visualization of a few TIFF images, that have not been preprocessed. Figure 2 shows two sample image slices of the dimension 512×512 from training dataset accompanied by the plot of their respective masks. Purple indicates the region(s) not containing glomeruli and yellow indicates the region(s) containing glomeruli. The HuBMAP test dataset consists of 5 public test images, stored in the TIFF file format. Just like training images, corresponding mask for them has also been provided.

4.2. Exploratory Data Analysis

Our dataset contains ample of details around each image sample, useful for exploratory data analysis in order to evaluate the goodness of data with respect to diversity, correctness, and usability. The fields in this dataset are the patient's race, ethnicity, sex, age, weight in kilograms, height in centimeters, bmi in kg/m^2, followed by kidney laterality, percent of area covered by cortex, and percent of area covered by medulla for each image sample.

We ran a thorough exploratory data analysis on the training and the test data in order to discover the trends and patterns in them.

The training data consists of 15 unique patients and test data consists of 5 unique patients.

The training data has a fair mix of samples from patients from different races(White and Black/African American) while the test data has samples only from White patients. The data does not contain any samples from patients of Hispanic or Latino ethnicity.

The training set contains of comparable number of samples from males and females but the test set contains only 1 sample from males. Hence, there is considerable skew in the test data in this aspect.

The age range of patients in the data also seems to be extremely limited and inclined towards aged patients. All the patients in both the datasets are aged between 58 and 76. However, the distribution of samples across this age range is almost uniform in both the training and test datasets.

Coming to the laterality, we have a few more samples from the right kidney in both the datasets.

Through exploratory data analysis, we could also verify that the combined area of medulla and cortex as reported in the dataset for each image sums to 100. Further, all kidney images in the dataset have more region covered by the cortex and not so much by the medulla. This is helpful because glomeruli units are found inside the cortex and not the medulla, thus our dataset has concentrated information about the glomeruli.

We also ran data analyses to find trends using a combination of parameters- gender and age, race and age, race and medulla region, and so on. The results from all these exploration tracks were nearly uniform as expected and desired. No alarming skews detected.

Finally, we drew a correlation matrix(figure 3) among the parameters of the data. Through this, we were also able to run certain sanity checks on the data and derive certain unknown correlations.

A very high correlation was found in the BMI and the weight of patients, which is natural and expected.

Considerable correlation was also found between weight and height of patients, although not as high as a layman would expect.

Percentage of cortex and percentage of medulla were found

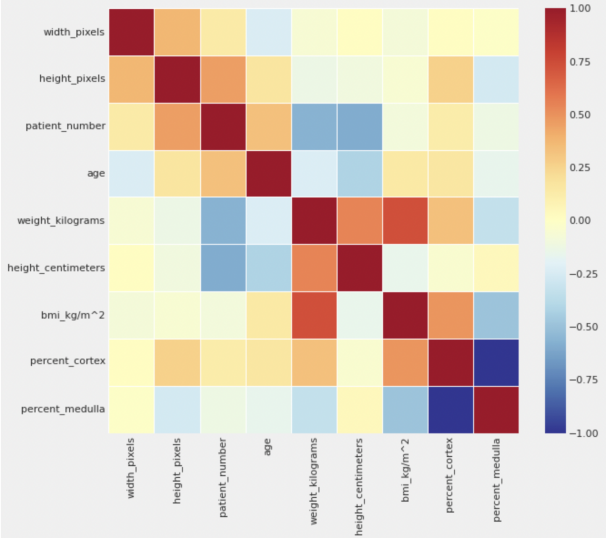

Figure 3: Correlation matrix on features of training samples

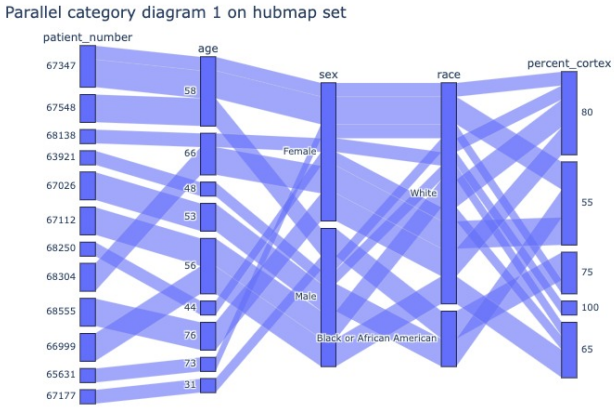

Figure 4: Parallel category diagram illustrating the trend of features for each image sample in training set

to be negatively correlated, which further validates our theory of total area occupancy.
Interestingly, patient's age and weight were found to have low negative correlation while age and BMI had low positive correlation.
Overall, the training and test data were found to be suitable for our application with satisfactory distribution across age, genders, and expected mathematical correlations.

4.3. U-Net Implementations

4.3.1 Epoch Count

All the experiments on U-Net have been on run for 15 epochs for a fair comparison between the accuracy and loss metrics of models trained using different learning rates.

4.3.2 Loss Function

The Dice Loss function is used in our U-Net implementation which was described in the previous section. For every epoch, training loss and validation loss were calculated. Finally, test loss was calculated using the best version of the model.

4.3.3 Baseline Model

We first implemented a baseline model publicly available on Kaggle. It is described below.

Table 1: Baseline model description

Network	U-Net
Training Loss	Dice Loss
Optimizer	Adam
Learning Rate	1e-5

The baseline model trains for 15 epochs and its results are summarized in below.

Table 2: Baseline model metrics

	Dice Loss	Accuracy
Training Data	0.1810	96.84%
Test Data	0.2399	93.46%

4.3.4 Other Models

We experimented with more models using the U-Net network, and successful results from the same are illustrated below. The experimental models were implemented by changing the learning rates and top 3 metrics were obtained from the following models.

Table 3: Model 1(learning rate = 4e-6) metrics

	Dice Loss	Accuracy
Training Data	0.415	88.93%
Test Data	0.472	85.84%

Table 4: Model 2(learning rate = 5e-5) metrics

	Dice Loss	Accuracy
Training Data	0.124	97.85%
Test Data	0.175	94.73%

Table 5: Model 3(learning rate = 5e-4) metrics

	Dice Loss	Accuracy
Training Data	0.1508	94.48%
Test Data	0.321	92.61%

4.4. ResNet50 Implementations

We also used the ResNet50 model on this task of classifying pixels of the given kidney images into 'glomerulus'and 'not glomerulus'.

4.4.1 Epoch Count

The baseline model almost fully converges to a good value of Dice Loss at the end of 15 epochs, hence all our ResNet50 implementations have been run for 15 epochs in order to directly compare the results.

4.4.2 Loss Functions

The following two loss functions are significant to our ResNet50 implementation

- Dice Loss
- Weighted loss = 0.8*BCE Loss + 0.2*Dice Loss

The BCE Loss here refers to the Binary Cross Entropy loss between the calculated segmentation and the target segmentation and Dice Loss has been defined in previous sections. The weighted loss has been introduced for the purpose of backpropagation while training.

4.4.3 Pretrained Coco Weights

All our ResNet50 implementations start with pretrained coco weights in order to aid faster training of the networks. Hence, the loss metric of most our models are quite low even as we start the training. Using the pretrained weights for our model initialization also helps the models in converging quickly.

4.4.4 Baseline Model

We first implemented a baseline model that is publicly available on Kaggle. It is described below.

Table 6: Baseline model description

Network	ResNet50
Training Loss	Weighted Loss
Optimizer	AdamW(weight decay = 1e-3)
Learning Rate	1e-4

The baseline model trains for 15 epochs and its results are summarized below.

Table 7: Baseline model metrics

	Dice Loss	Accuracy
Training Data	0.107	98.03%
Test Data	0.162	94.94%

4.4.5 Other Models

The following model was implemented by merely changing the baseline model's optimizer from AdamW to Adam, as the most initial experiment. Metrics of this model were found to be the worst among all implemented models.

Table 8: Model 1 description

Network	ResNet50
Training Loss	Weighted Loss
Optimizer	Adam
Learning Rate	1e-4

Table 9: Model 1 metrics

	Dice Loss	Accuracy
Training Data	0.472	86.01%
Test Data	0.415	88.12%

We then implemented more models using the ResNet50 network, and the notable results from those are illustrated below. The next model was implemented by using Dice Loss as our training loss instead of the Weighted Loss. The metrics of this model are not just comparable to the baseline model, but slightly better in performance on the test data.

Table 10: Model 2 description

Network	ResNet50
Training Loss	Dice Loss
Optimizer	AdamW(weight decay = 1e-3)
Learning Rate	1e-4

Table 11: Model 2 metrics

	Dice Loss	Accuracy
Training Data	0.108	98.01%
Test Data	0.157	95.00%

The following model was implemented by using Dice Loss as the training loss and Adam optimizer. The performance metrics of this model are good but certainly not the best.

Table 12: Model 3 description

Network	ResNet50
Training Loss	Dice Loss
Optimizer	AdamW(weight decay = 1e-3)
Learning Rate	1e-4

Table 13: Model 3 metrics

	Dice Loss	Accuracy
Training Data	0.154	95.03%
Test Data	0.194	93.88%

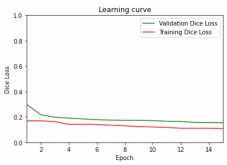

Figure 5: Learning curve of the best ResNet50 model

The following model was implemented by using Weighted Loss as the training loss, Adam as the optimizer, and increasing learning rate to 1e-3.

Table 14: Model 4 description

Network	ResNet50
Training Loss	Weighted Loss
Optimizer	Adam
Learning Rate	1e-3

Table 15: Model 4 metrics

	Dice Loss	Accuracy
Training Data	0.169	94.01%
Test Data	0.181	93.70%

The following model was implemented by using Weighted Loss as the training loss, AdamW as the optimizer, and 1e-3 as the learning rate.

Table 16: Model 5 description

Network	ResNet50
Training Loss	Weighted Loss
Optimizer	AdamW(weight decay = 1e-3)
Learning Rate	1e-3

Table 17: Model 5 metrics

	Dice Loss	Accuracy
Training Data	0.150	95.20%
Test Data	0.190	94.38%

5. Conclusion

The paper focuses on detecting glomeruli in kidneys using image segmentation techniques such as U-Net and ResNet. After training and testing the proposed algorithms, the best accuracy for U-Net was found to be 94.73% with a dice coefficient of 0.825 and the best accuracy for ResNet was found out to be 95% with a dice coefficient of 0.843. Both of these models seem very promising as even the worst accuracy values in our experiments were 88.12% and 85.84% respectively for ResNet and U-net. As U-Net and ResNet are advanced architectures of convolutional networks, in future work other neural networks that utilize CNN layers can be explored with more creative combinations of optimizers and learning rates.

References

[1] S. L. Vaden, "Glomerular disease," *Topics in companion animal medicine*, vol. 26, no. 3, pp. 128–134, 2011. 1

[2] B. de Bono, P. Grenon, R. Baldock, and P. Hunter, "Functional tissue units and their primary tissue motifs in multi-scale physiology," *Journal of biomedical semantics*, vol. 4, no. 1, pp. 1–13, 2013. 1

[3] N. Bouteldja, B. M. Klinkhammer, R. D. Bülow, P. Droste, S. W. Otten, S. F. von Stillfried, J. Moellmann, S. M. Sheehan, R. Korstanje, S. Menzel *et al.*, "Deep learning–based segmentation and quantification in experimental kidney histopathology," *Journal of the American Society of Nephrology*, vol. 32, no. 1, pp. 52–68, 2021. 2

[4] C. P. Jayapandian, Y. Chen, A. R. Janowczyk, M. B. Palmer, C. A. Cassol, M. Sekulic, J. B. Hodgin, J. Zee, S. M. Hewitt, J. O'Toole *et al.*, "Development and evaluation of deep learning–based segmentation of histologic structures in the kidney cortex with multiple histologic stains," *Kidney international*, vol. 99, no. 1, pp. 86–101, 2021. 2

[5] G. Bueno, M. M. Fernandez-Carrobles, L. Gonzalez-Lopez, and O. Deniz, "Glomerulosclerosis identification in whole slide images using semantic segmentation," *Computer methods and programs in biomedicine*, vol. 184, p. 105273, 2020. 2

[6] C. Zeng, Y. Nan, F. Xu, Q. Lei, F. Li, T. Chen, S. Liang, X. Hou, B. Lv, D. Liang *et al.*, "Identification of glomerular lesions and intrinsic glomerular cell types in kidney diseases via deep learning," *The Journal of pathology*, vol. 252, no. 1, pp. 53–64, 2020. 2

[7] E. Uchino, K. Suzuki, N. Sato, R. Kojima, Y. Tamada, S. Hiragi, H. Yokoi, N. Yugami, S. Minamiguchi, H. Haga *et al.*, "Classification of glomerular pathological findings using deep learning and nephrologist–ai collective intelligence approach," *International Journal of Medical Informatics*, vol. 141, p. 104231, 2020. 2

[8] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer vision, graphics, and image processing*, vol. 29, no. 1, pp. 100–132, 1985. 2

[9] R. L. Barkau, "Unet: One-dimensional unsteady flow through a full network of open channels. user's manual," Hydrologic Engineering Center Davis CA, Tech. Rep., 1996. 2

[10] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440. 2

[11] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014. 3

[12] Z. Zhang, "Improved adam optimizer for deep neural networks," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*. IEEE, 2018, pp. 1–2. 3

[13] J. F. Muhammad Yaqub and M. S. Zia, "State-of-the-art cnn optimizer for brain tumor segmentation in magnetic resonance images," *Brain sciences vol. 10,7 427*, 2020. 3

[14] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv:1711.05101*, 2019. 3

[15] H. Consortium *et al.*, "The human body at cellular resolution: the nih human biomolecular atlas program," *Nature*, vol. 574, no. 7777, p. 187, 2019. 4