Diabetic Complications Consortium

Application Title: Automated quantification of ultrastructural pathology of diabetic nephropathy using deep learning.

Principal Investigator: Behzad Najafian, M.D.

1. Project Accomplishments:

We have developed a novel validated deep learning model for foot process width (FPW) measurement on electron microscopy (EM) images.

2. Specific Aims:

Aim #1: To develop deep learning models for accurate estimation of foot process width in diabetic nephropathy.

Results: We developed a novel model named Forknet for foot process width measurement. Our accomplishments are summarized below:

1. Model Design: A custom U-Net based model was designed to detect both PGBMI and the podocyte filtration slit diaphragms in two separate layers (Figure 1).
2. Post process script for FPW calculation: The calculation of the FPW was done in a post process script.
3. Model training: In order to reduce over-fitting and help improve validation accuracy during training, the images went through a randomized augmentation process whereby they were randomly scaled, flipped, shifted (adjusting mean and contrast), and rotated. The validation dataset was not augmented to keep validation scores closer to their real values and evaluate the training and test accuracy divergence. Images, during the training phase, were manually segmented with a 10% test split, meaning that every 1 in 10 images was not included in the model optimizer to allow detection of model under-fitting or model overfitting. Training involved a semi-online learning method, which initially trained the model on a small sample. When the manual segmentation data count reached ~200 images, the model was used in the Segmentation Utility to segment new images, which were corrected by our team to be appended to the dataset. The training software reanalyzed the entire dataset and reduced the learning rate of the model as the accuracy plateaued. Only the highest accuracy, measured by dice similarity coefficient (DSC), models were saved for later use. This process was continued until the training dataset was complete. Once complete, the entire dataset was run to capture the highest validation DSC. Once a plateau was reached and the learning rate was dropped to 0, the model weights were repackaged with the Report Utility, and made available for remote predictions.

4. Compiling the package: After successfully trained on the entire dataset, the user interface, pre-processing, post-processing, weights, and architecture, and vision/deep learning frameworks were all compiled into one application.
5. Report Utility: We built a post-processing tool called UW Report Utility that can process and view the predictions from the network in real time. The workflow and the post-processing steps are shown in Figure 2.
6. Model validation: Performance of the model was analyzed using two metrics, DSC and model accuracy. Each manual segmentation mask, per category, was compared to the prediction mask of the model. DSC was calculated by dividing the intersection by the total area of the masks, with an additional smoothing constant of 1 DSC = $(2|X \cap Y| + 1)/(|X| + |Y| + 1)$ for each category. DSC values range from 0-1, where 0 represents no pixel overlap and 1 represents perfect pixel overlap between the two masks. DSC values obtained from testing different set are provided in Table 1.
7. Testing the model on diabetic biopsies. FPW was measured on digital electron microscopy images (~30,000x) obtained using systematic uniform random sampling from type 1 diabetic patients. The results are provided in table 2.

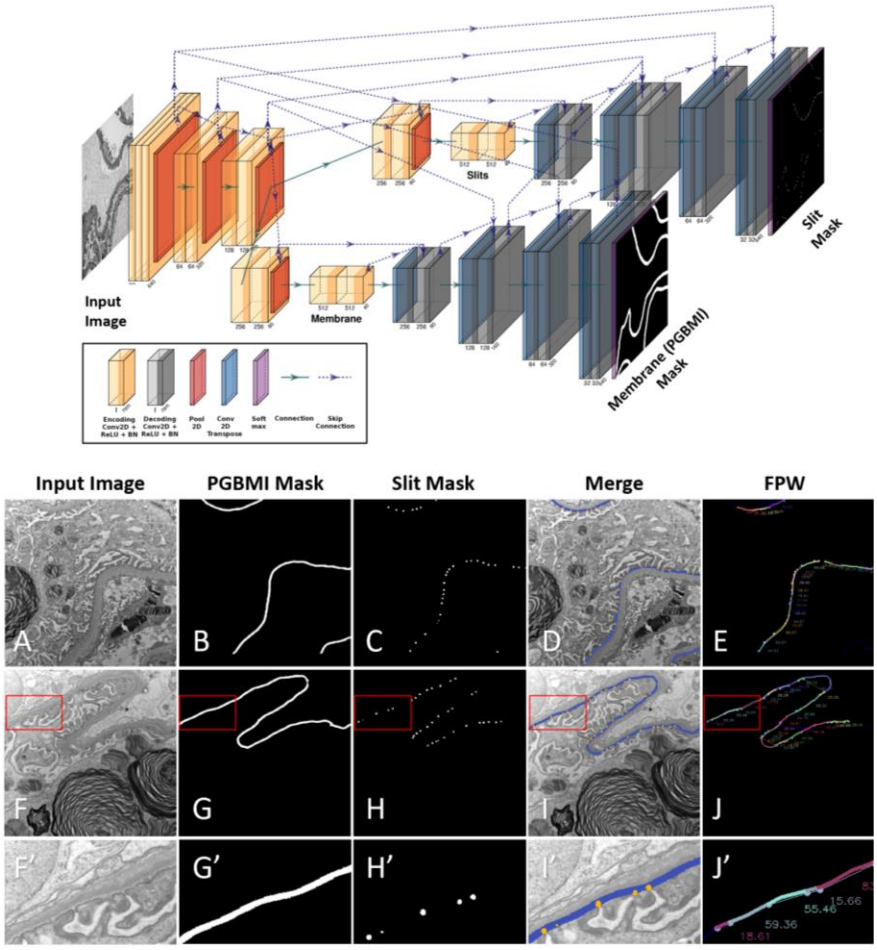

Figure 1. Model of the ForkNet Architecture. Top: A schematic representation of the architecture showing an input EM Image to the left, conv2d blocks max pooling for downscaling, split branches, deconv2d and conv2d blocks for upscaling, and output layers to the right. Layer

connections are identified as solid directional lines, and skip connections are identified as dashed directional lines. Note: the skip connection from the membrane branch to the slit branch. Most residual connections used convolutional blocks rather than identity connections. Bottom: A-E and F-J show two examples of input images (A and F), resulting in corresponding PGBMI (B and G) and slit masks (C and H) with merged images, showing the masks superimposed on the input images (D and I); and the post-processing output with individual foot process width (FPW) measurements (E and J). F'-J' show magnified views of the red boxes in corresponding F-J images. The size of the dots in the slit mask and merged images reflects the model confidence in prediction of a slit.

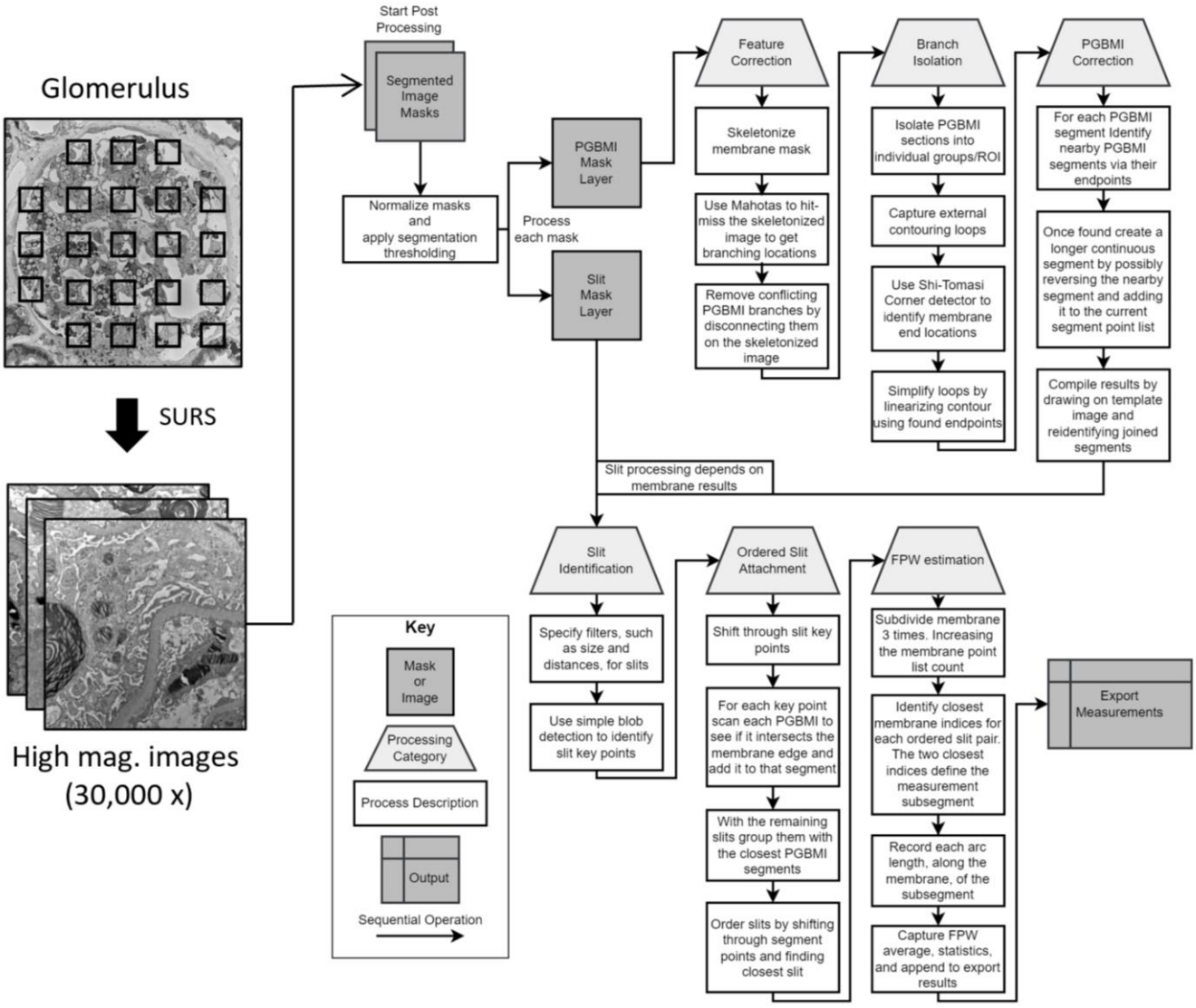

Figure 2. ForkNet post-processing script general workflow. Top left image: a low magnification view of a glomerulus. The small squares represent systematic uniform random sampling (SURS) images taken at higher magnification (30,000 x) shown in bottom left that are used for foot process width (FPW) measurements. The flowchart on the right shows the post-processing workflow. The input images are normalized and go through various filters and processing, the

output of which will be semantic segmentation of filtration slits and podocyte – glomerular basement membrane interface (PGBMI). The length of the individual segments (PGBMI limited between two adjacent slits) is measured and exported. B. ForkNet model architecture. On the left, an input EM Image is shown (30,000 x, obtained using SURS). Different components of the architecture are shown in color coded boxes that are connected through solid or dashed lines as explained in the box at the bottom left corner. In summary, from left to right, the input image goes through conv2d downscaling blocks, split branches, and deconv2d upscaling to result in PGBMI and slit output masks.

Table 1. Examples of dice coefficient (DSC) values for various training groups for podocyte - glomerular basement membrane interface (PGBMI) and filtration slit masks used for evaluation of model performance

Training Group	Mean DSC for PGBMI	Mean DSC for Slit
1	0.76	0.56
2	0.36	0.27
3	0.72	0.48
4	0.70	0.56
5	0.78	0.58
6	0.72	0.46
7	0.74	0.48

Table 2. Output data from research biopsies from patients with type 1 diabetes:

	File Count	Average Time	Average FPW	STD FPW	Membrane Length	Average Attachment	STD Attachment
1	9.356783867	0.145738572	88.75298309	69.90515137	178957.4996	350.56427	274.1158447
2	13.66515827	0.160995133	58.45458221	29.50606918	193869.9907	383.7431946	292.4305725
3	4	0.093974691	93.46107483	80.22080994	134298.8327	215.7985992	248.3533478
4	11.8197279	0.160650975	72.82346344	35.89996338	294852.5307	386.850708	275.3671265
5	7.361022472	0.125372871	76.33524323	44.09388351	231365.8618	286.0407715	267.0921021
6	12.60063934	0.159464335	67.53728485	32.01740646	292522.6741	424.4464722	298.3026428
7	5.98507452	0.104407704	77.49528503	46.45928955	76315.63316	282.505127	278.6502075
8	14.1459856	0.182454294	71.86685944	37.06433105	152762.2283	473.7096558	329.6286621
9	9.08791256	0.128895912	69.29936218	44.42921829	124206.1295	337.879364	377.1394653
10	19.21904755	0.217414604	60.20162964	26.65054512	259508.7894	456.6531677	308.2755737
11	17.52962875	0.202325484	58.04418564	27.12368393	308206.4905	423.4949036	322.8369141
12	6.569832325	0.119271859	94.48657227	86.40921783	132805.7282	255.3275909	277.9932251
13	6.344444275	0.114430881	75.62589264	44.28681183	118194.1071	269.6115112	263.5831299
14	6.857142925	0.120465058	96.43178558	80.59918213	164940.8217	253.5434875	240.4208527

Aim #2: To develop DL models for accurate semantic segmentation of GBM and accurate estimation of GBM width in DN.

1. Developing a Segmentation Utility: Creating a custom multi-class utility called UW Segmentation Utility using wxPython to manually classify podocyte-glomerular basement membrane interface (PGBMI) and foot process slits (Figure 3).
2. Model training: A total of 8,152,303 nm of PGBMI were manually segmented in the training set.
3. Skeletonization of the GBM (ie get the center line) and measuring the GBM width at random points. At each point we take the tangent of the line and then move a step away from the center defined as dx, at this new step we get the new tangent (or line distance that's closest to the membrane center) and move in that direction and keep going until we hit the edge of the membrane (Figure 4). We are still working on this part.

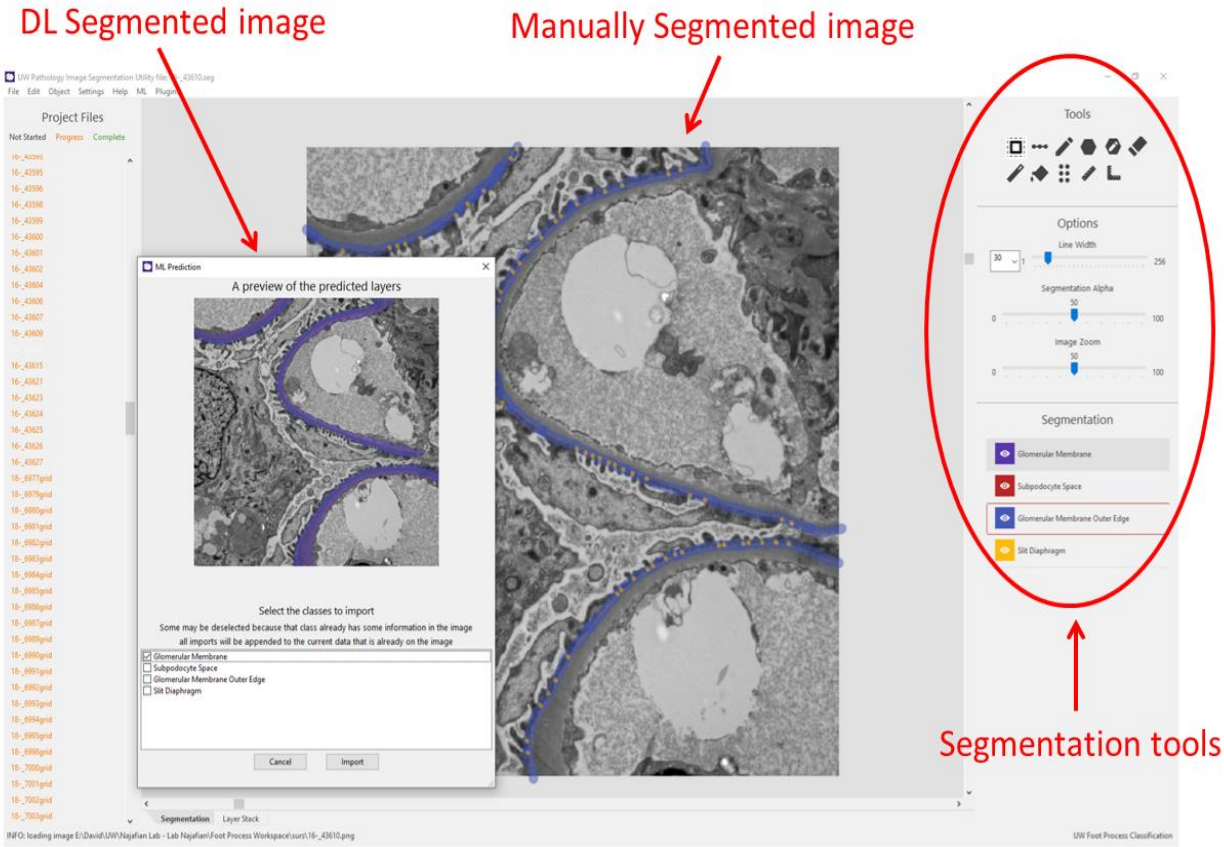

Figure 3. A screenshot from UW Segmentation Utility. The column on the left shows the list of images (orange). The smaller image shows a preview of the predicted layers by deep learning and the larger image shows the manually segmented or corrected (based on the model prediction) image which will be exported to training set. On the right, the segmentation tools are seen.

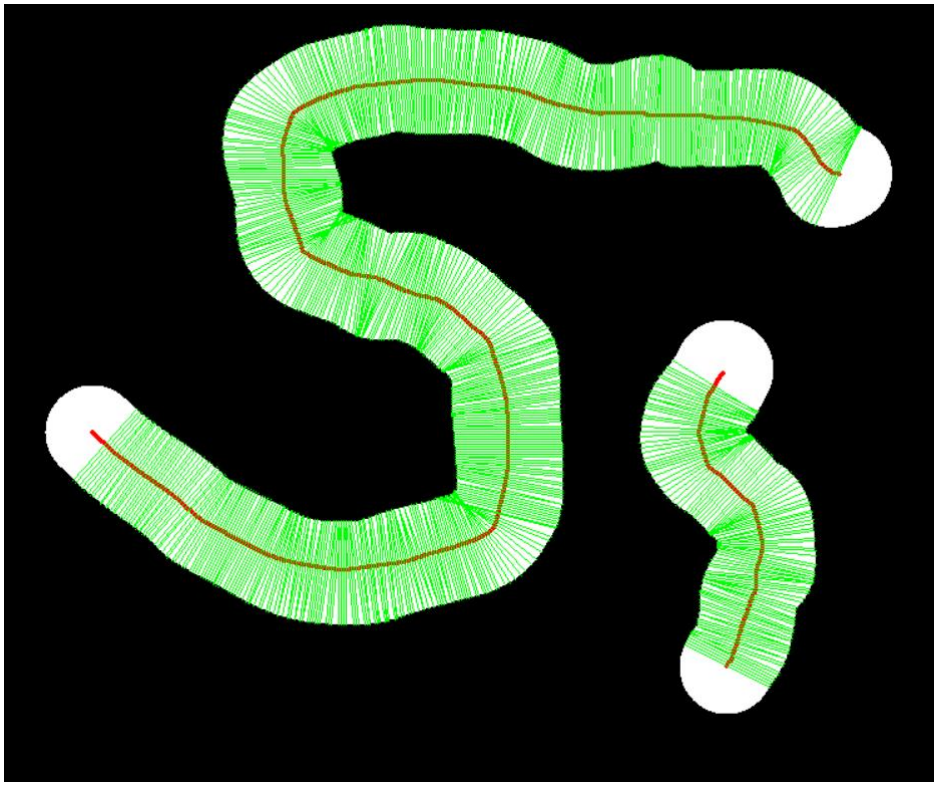

Figure 4. Glomerular basement membrane is skletonized (red middle line) and its width is measured at many random points (green lines)

Aim #3: To develop DL models for accurate semantic segmentation and quantification of mesangium and mesangial matrix in DN.

1. The same segmentation utility developed for Aim#2 is being used for segmentation of mesangial cells and mesangial matrix and preparation of a training set. This part is still under development and refining.
2. In order to better understand the structure and morphology of mesangium and mesangial matrix we have been making 3D models of mesangium using serial block face scanning electron microscopy using a combination of Amira and Aivia softwares (Figure 5).

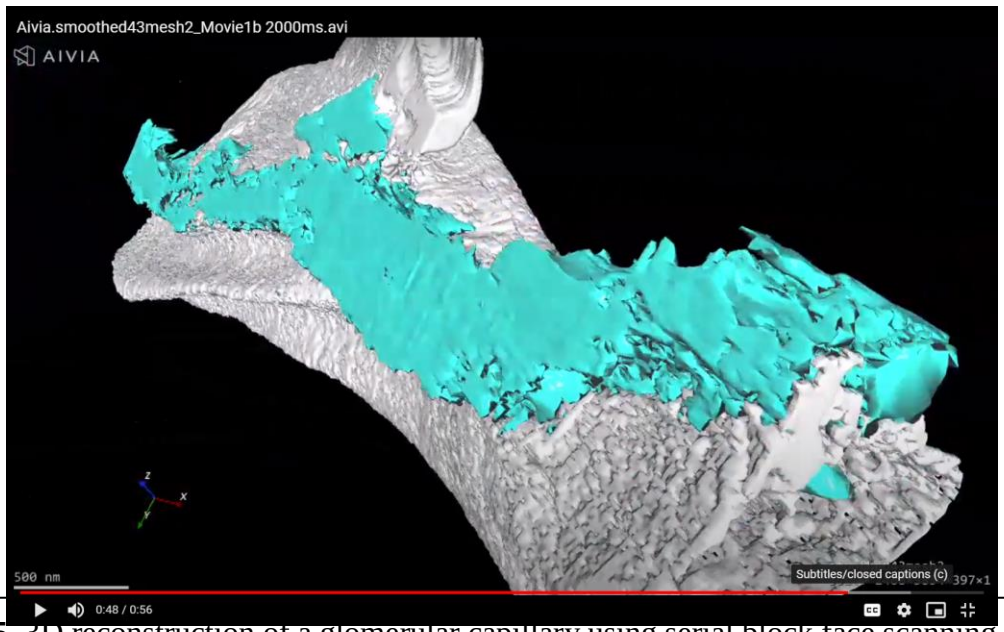

Figure 5. 3D reconstruction of a glomerular capillary using serial block face scanning electron microscopy. The light grey color marks the capillary and the complex blue structure is a mesangial cell covering the capillary.

3. Publications:

A manuscript describing the FPW measurement by deep learning is in preparation.