

Article

A Self-Adaptive Strip Pooling Network for Segmenting the Kidney Glomerular Basement Membrane

Caifang Song ^{1,*}, Xiangsheng Huang ² and Xiangyu Lyu ¹ 

¹ School of Mathematics and Computer Science, Shanxi Normal University, Taiyuan 030031, China; lvxiangyu11@gmail.com

² Hebei Key Laboratory of Cognitive Intelligence, Xiong'an Institute of Innovation, Xiong'an New Area 071700, China; huangxiangsheng@xii.ac.cn

* Correspondence: songcf@sxnu.edu.cn; Tel.: +86-1513-517-6061

Abstract: Accurate semantic segmentation and automatic thickness measurement of the glomerular basement membrane (GBM) can aid pathologists in carrying out subsequent pathological diagnoses. The GBM has a complex ultrastructure and irregular shape, which makes it difficult to segment accurately. We found that the shape of the GBM is striped, so we proposed an RSP model to extract both the strip and square features of the GBM. Additionally, grayscale images of the GBM are similar to those of surrounding tissues, and the contrast is low. We added an edge attention mechanism to further improve the quality of segmentation. Moreover, we revised the pixel-level loss function to consider the tissues around the GBM and locate the GBM as a doctor would, i.e., by using the tissues as the reference object. Ablation experiments with each module showed that SSPNet can better segment the GBM. The proposed method was also compared with the existing medical semantic segmentation model. The experimental results showed that the proposed method can obtain high-precision segmentation results for the GBM and completely segment the target. Finally, the thickness of the GBM was calculated using a skeleton extraction method to provide quantitative data for expert diagnosis.

Keywords: self-adaptive strip pooling; attention mechanism; glomerular basement membrane (GBM); skeleton extraction



Academic Editor: Steve Ling

Received: 25 December 2024

Revised: 8 February 2025

Accepted: 10 February 2025

Published: 14 March 2025

Citation: Song, C.; Huang, X.; Lyu, X. A Self-Adaptive Strip Pooling Network for Segmenting the Kidney Glomerular Basement Membrane. *Sensors* **2025**, *25*, 1829. <https://doi.org/10.3390/s25061829>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In 2002, the National Kidney Foundation (NKF) formulated a concise definition of chronic kidney disease (CKD) and standardized the diagnostic criteria, thereby facilitating extensive epidemiological studies targeting the general population. This move enabled researchers to gain a more comprehensive understanding of the prevalence, progression, and associated factors of CKD [1]. A large number of studies on people with kidney disease have been carried out around the world. The results of these studies showed that the incidence of kidney disease is approximately 10%, while only approximately 12% of these people know that they are sick, and that an imbalance between the number of relevant pathologists and patients leads to inefficient diagnoses [2]. CKD has become an important disease that threatens global public health. Kidney puncture biopsy pathology is an important means for diagnosing kidney disease. Transmission electron microscopy is an effective technique for CKD screening. Using the advanced technology of transmission electron microscopy, healthcare professionals can gain a deeper understanding of the intricate, submicroscopic changes that occur within glomerular cells in pathological conditions. This heightened level of detail allows for a more thorough and accurate pathological diagnosis,

empowering doctors to make more informed treatment decisions and ultimately improve patient outcomes.

It can be observed in Figure 1 that the gray image texture corresponding to the glomerular ultrastructure is complex, so it is difficult to identify and judge pathological changes. It is very difficult to observe such complicated pathological images and make diagnoses with the naked eye for an extended period. In the submicroscopic structure of glomerular cells, changes in the glomerular basement membrane (GBM) are closely related to chronic kidney disease (CKD). Doctors often need to identify and measure the GBM in the process of pathological diagnosis. Most of the GBM has low contrast with the surrounding tissue structures. It is not only difficult but also time-consuming to identify changes with the naked eye. Employing computer technology to aid pathologists in the precise and automated identification of the GBM holds immense potential for mitigating diagnostic challenges and enhancing both the efficiency and precision of pathological assessments. This technological advancement stands to revolutionize the field by streamlining the diagnostic process and ensuring a higher degree of accuracy, ultimately benefiting patients through more informed treatment decisions.

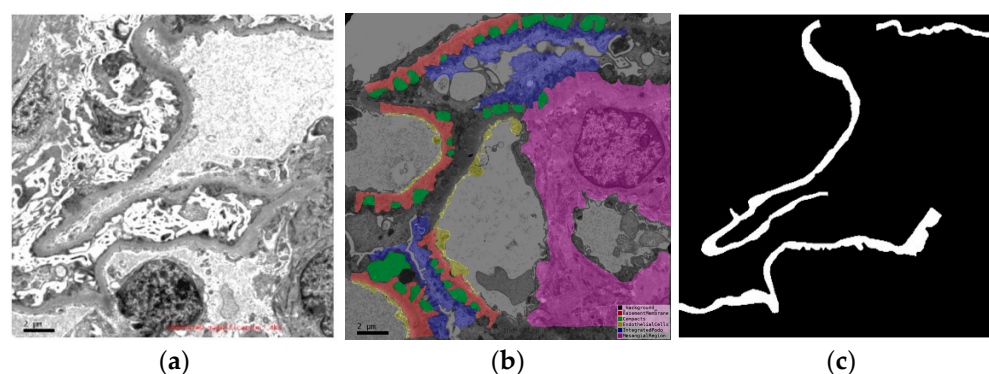


Figure 1. An electron microscopic image derived from a renal biopsy. (a) The original electron microscopic image. (b) A comprehensively annotated version, where various components are highlighted in distinct colors: the glomerular basement membrane (GBM) in red, compacts in green, endothelial cells in yellow, the mesangial region in blue, and podocytes in purple. (c) The GBM, with the image showcasing a refined marking of this critical structure alone.

In recent years, deep learning techniques have undergone remarkable advancements in the domain of multi-structural medical image segmentation, marking a significant milestone in enhancing the precision and efficiency of this complex task. In this paper, based on the striping features of the GBM, strip pooling and standard pooling were used to extract the texture of the GBM. At the same time, the grayscale of the GBM is similar to that of surrounding tissues, thus leading to low contrast. To further elevate the quality of segmentation outcomes, this paper incorporated an edge attention mechanism, a strategic enhancement aimed at refining the delineation of structural boundaries within medical images. Finally, we revised the pixel-level loss function to consider tissues around the GBM and located the GBM like a doctor would by taking the tissue as the reference object.

2. Related Work

Early studies on the automatic semantic segmentation of the GBM were focused on traditional, artificially designed grayscale, morphological, and textural features. As far back as 1993, Ong et al. [3] introduced the groundbreaking use of self-adaptive window-based tracking techniques for segmenting glomerular electron microscopy images, setting a precedent in the field. From then on, some semi- and full-automatic methods have been proposed. By employing the advanced method of region segmentation coupled with dynamic contour

modeling, Kamenesky [4] and Rangayyan et al. [5] were able to effectively segment and quantify the GBM, achieving precise measurements and insights. Wu et al. [6,7] presented two methods: The first involves obtaining the centerline of the GBM by interpolating manually marked points, followed by automatic segmentation of the GBM using distance mapping and low-pass filtering techniques. The other method uses a threshold method and a morphological method without the involvement of artificial markers. In addition to segmenting the GBM, Liu et al. [8] also measured its length and calculated the number of gaps present. In the past two years, new segmentation methods, such as image block matching [9] and the use of particle filters [10], have been gradually introduced. In 2017, Guo et al. [11] introduced a novel algorithm based on neutrosophic sets, and shearlet transform has been proposed to enhance the accuracy of glomerular basement membrane segmentation. This approach effectively improves segmentation precision and efficiency by integrating features extracted through the shearlet transform with neutrosophic image processing. In 2019, Cao et al. used random forests to segment the GBM. This method achieved good results, but it is not good enough for medical image segmentation because the false positive rate is high [12]. While the aforementioned methods have indeed contributed to the field, numerous challenges remain unaddressed. These approaches often rely heavily on manual initialization, which not only imposes an additional burden on pathologists but also introduces potential subjective biases. Additionally, they are typically limited to segmenting truncated GBM fragments, focusing primarily on contrast enhancement and unidirectionality. Consequently, ensuring high-quality segmentation of complex GBM images in their entirety remains an ongoing challenge.

In recent years, multi-structure medical image segmentation has witnessed significant advancements driven by deep learning methods. Lei et al. conducted a comprehensive survey exploring the various applications of deep learning techniques in medical image segmentation [13,14]. Most existing medical images are clustered, such as images of the pancreas, retina, brain tumors, and bone marrow. However, the GBM is striped. Current techniques employed for semantic segmentation in medical imaging are not readily transferable to the specific task of GBM segmentation, necessitating alternative or adapted strategies. In 2020, Qi bin Hou et al. [15] introduced a strip pooling approach that markedly improves the capture of long-range dependencies. We further refined this method and integrated it into existing medical image semantic segmentation frameworks, thereby enhancing the segmentation of intricate strip texture features. The new model, called self-adaptive strip pooling (SSP), automatically extracts both the strip and square features of the GBM.

3. SSPNet Method

Semantic segmentation models are usually encoding–decoding frameworks. The input of the encoder is downsampled to concentrate information, while the decoder is responsible for upsampling to restore the input size. In the context of semantic segmentation, the encoder plays a pivotal role in extracting crucial contextual information about objects, necessitating a robust multiscale capability. The Res2Net framework achieves this by incorporating hierarchical residual-like connections within a single residual block. This approach enables the model to represent features across multiple scales at a granular level, thereby enhancing the receptive field range for each network layer [16]. Therefore, we used Res2Net as the base model in this paper. To elaborate further, for a given input image, our approach involves extracting five distinct levels of features, denoted as f_i , where i ranges from 1 to 5, utilizing a Res2Net-based backbone network. Subsequently, we categorized these features into two groups: low-level features, comprising f_1 and f_2 , and high-level features, which encompass f_3 , f_4 , and f_5 .

For the decoder, a loss of information due to downsampling is crucial for segmentation tasks. A bilinear upsampling process usually fails to restore information loss. To further enhance the Res2Net network, we integrated dilated convolutions, facilitating exponential expansion of the receptive field while maintaining resolution and coverage. Leveraging the strengths of Res2Net, we introduced a novel deep neural network architecture, called the self-adaptive strip pooling network (SSPNet), for the GBM segmentation task. As shown in Figure 2, the sizes of f_3 , f_4 , and f_5 are the same because of the use of dilated convolution.

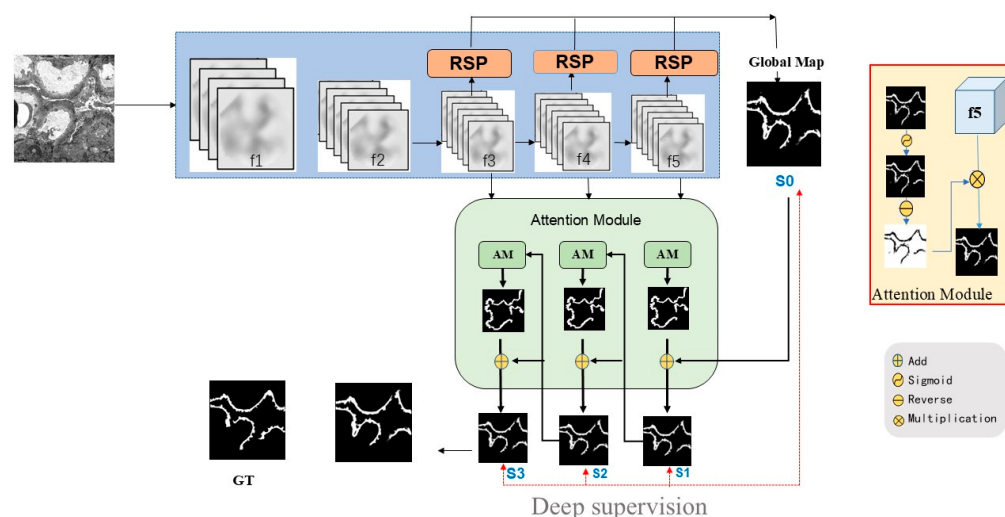


Figure 2. Overview of the proposed SSPNet. The backbone is Res2Net. Due to dilated convolution, f_3 , f_4 , and f_5 are the same size. We added an RSP module to extract both strip and square features. A set of attention modules were added to enhance the details. We employed deep supervision for three side outputs (S_1 , S_2 , S_3) and the global map (S_0), as indicated by the red dotted line, to ensure optimal performance across all stages of the model.

First, we observed that the shape of the GBM exhibits a striped pattern. Recognizing the limitation of traditional pooling methods in adequately capturing such strip information, we devised and incorporated a refined strip pooling (RSP) module into our framework. This RSP module was specifically designed to extract and integrate the mixed information embedded within the striped features, thereby enhancing the overall feature representation capabilities of our model.

Second, grayscale images of the GBM are similar to grayscale images of surrounding tissues, and the contrast is low. Therefore, an attention module was added to strengthen the edge information.

Finally, in order to mimic the expertise of doctors in locating the GBM while considering the surrounding tissue, we introduced a novel loss function. This function treats the tissue as a reference object, ensuring that the GBM is identified in a manner similar to that of a medical professional approaching the task.

We integrated attention mechanisms and a novel loss function. These mechanisms enable the network to automatically learn and adjust the importance of different features, thereby achieving self-adaptation and optimization of feature representations. Consequently, our network was named the self-adaptive strip pooling network (SSPNet). The details of each component are explained in the following sections.

3.1. Self-Adaptive Strip Pooling Module

While the semantic segmentation method, rooted in the fully convolutional network (FCN), has proven adept at capturing high-level semantic information, it predominantly relies on local convolution and pooling operations, resulting in a constrained receptive

field. This limitation hampers its effectiveness in analyzing complex scenes, leaving room for improvement. Researchers may use attention mechanisms to model long-distance dependencies [17,18]. Alternative strategies to address the issue of a limited receptive field in semantic segmentation methods include utilizing dilated convolutions [19–21], which broaden the receptive field without introducing additional parameters, or employing global or pyramid pooling techniques [22] to enrich global contextual cues. However, the limitation of these methods is that they are mainly used when the input feature is square, and the effect is not very good for many long-strip targets.

In certain scenarios, the object of interest may exhibit a lengthy strip-like structure, as exemplified by the GBM depicted in Figure 1. Using a large square pooled window does not capture bar information well. The strip pooling method can capture long dependencies more effectively [15].

Given a two-dimensional input tensor denoted as x , where x belongs to the space $R^{H \times W}$, H represents the spatial height, and W represents the spatial width. The size of the pooling window is $d \times d$, and the size of the output two-dimensional tensor Y is $H_0 \times W_0$. The standard pooling formula can be expressed as

$$y_{i_0 j_0} = \frac{1}{d \times d} \sum_{0 \leq i \leq d} \sum_{0 \leq j \leq d} x_{i_0 \times d + i, j_0 \times d + j} \quad (1)$$

where $0 < i_0 < H_0, 0 < j_0 < W_0, H_0 = H/d$, and $W_0 = W/d$.

It can be seen from Equation (1) that the standard pooling window is square, so it has an advantage in dividing square graphics. When the input image is striped, it cannot be segmented well.

The strip pooling method specifically solves the strip texture. Strip pooling differs from traditional two-dimensional average pooling in its unique window size, which is either $(H, 1)$ or $(1, W)$. Instead of averaging features across the entire two-dimensional plane, strip pooling focuses on averaging all features within a single row or column. Its expression is

$$y_i = \frac{1}{W} \sum_{0 \leq j \leq W} x_{i,j}, y_j = \frac{1}{H} \sum_{0 \leq i \leq H} x_{i,j} \quad (2)$$

where $y_i \in R^H$ and $y_j \in R^W$ respectively represent the strip pooling graph after strip pooling. Strip pooling enables the benchmark network to utilize long-distance dependencies effectively.

However, the strip pooling module (SPM) and the mixed pooling module (MPM) proposed in SPNet [15] have their own limitations. To effectively capture the varied influences posed by both long and non-long objects, we integrated both strip and pyramid pooling mechanisms into a composite mixed pooling module. This integration allows for a more comprehensive and nuanced treatment of feature extraction, catering to the distinct characteristics exhibited by diverse object shapes within the dataset. Therefore, we added both strip and pyramid pooling into a mixed pooling module, which considers the effects of long and non-long objects. However, the pyramid pooling operation is performed after the strip pooling operation, and there is no way to recover the information about the lost non-striped objects. The network is also very complicated. After carefully studying the entire network, the feature maps generated by the basic ResNet network at each stage are multiscale. In this paper, SSP operation was directly performed on the multiscale feature images f_3, f_4 , and f_5 generated by Res2Net, which not only makes the network model simpler but also takes into account the characteristics of strip and non-strip objects.

The RSP module is shown in Figure 3. This module employs horizontal and vertical strip pooling operations to gather contextual information along the spatial dimension,

specifically capturing horizontal and vertical features. After obtaining the features from the standard pooling and strip pooling techniques, we integrated them to arrive at the ultimate feature mapping. The refined strip pooling (RSP) module emphasizes the consolidation of diverse contextual information via diverse pooling methods, thereby enhancing the distinguishing capabilities of its features.

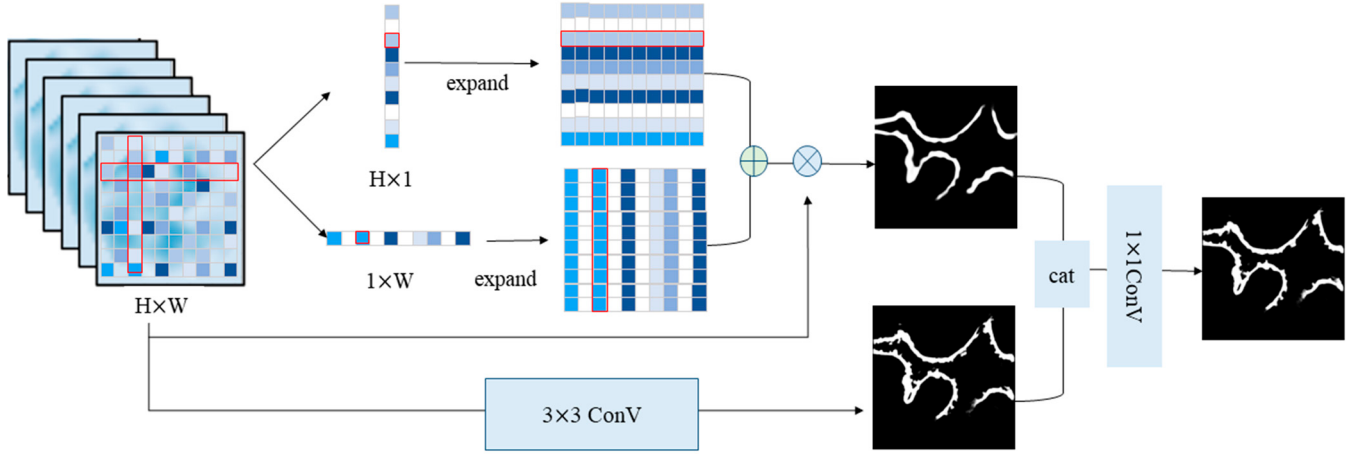


Figure 3. Schematic illustration of the refined strip pooling (RSP) module. This module incorporates both horizontal and vertical strip pooling methodologies to gather spatial context information along both the horizontal and vertical axes. It then connects the features obtained from the standard and strip pooling to obtain the final feature mapping. The red frames highlight the variation process of row or column vectors.

In addition to the strip-like characteristics of the GBM, it can also be observed from Figure 1 that the textures of the middle and edges of the GBM are different. This also creates challenges for the precise extraction of the GBM. In Figure 2, the global map generated by our network offers a coarse approximation of the GBM's location. To refine this, we employed a series of recurrent reverse attention modules that establish a correlation between regions and boundary cues, enabling a deeper exploration [23]. These attention modules were designed to dynamically learn from three parallel, sophisticated features. By applying reverse attention, our network is able to augment the boundary details of the GBM, resulting in a more precise localization.

Specifically, the feature denoted as the reversed attention output, R_i , is derived through an element-wise multiplication ($\odot$) operation between the high-level output features, $\{f_i$ for $i = 3, 4, 5\}$, and the corresponding reversed attention weights, R_i . This computation is formalized in the following equation:

$$R_i = f_i \odot A_i \quad (3)$$

The significance of the reverse attention weight, A_i , lies in its pivotal role in facilitating the detection of prominent objects within a given context, and it can be formulated as follows:

$$A_i = \ominus(\sigma(S_{i+1})) \quad (4)$$

where the sigmoid function is denoted by $\sigma(\cdot)$, and $\ominus(\cdot)$ represents a reverse operation that subtracts the input from a matrix E , where all entries are united. This erasing technique serves to enhance the prediction map, transforming crude and imprecise estimations into precise and comprehensive outcomes.

In Figure 2, the implementation of deep supervision is demonstrated for the three side outputs (namely, S1, S2, and S3) as well as the global map S0, indicated by the red dashed line. Each of these maps undergoes upsampling to ensure conformity with the dimensions of the ground truth map (GT). For each test, S3 achieved the best performance, so we chose S3 to compare with the GT image.

3.2. Loss Function and Deep Supervision

Due to the limited receptive field, low-level features retain an abundance of intricate details intertwined with background noise, which also have clear boundaries (critical to generating saliency maps). High-level features undergo numerous downsampling procedures, leading to a significant reduction in detailed information, but there are still consistent semantics and clear backgrounds. To focus more on difficult-to-classify samples in model training, self-adaptive loss is employed to diminish the influence of easily classifiable samples by decreasing their weights while simultaneously enhancing the emphasis on challenging-to-classify samples by increasing their respective weights. The loss function is defined as

$$\min L = L_{\text{IoU}}^{\omega} + L_{\text{BCE}}^{\omega} \quad (5)$$

where L_{IoU}^{ω} represents the weighted IoU. IoU is the result of dividing the overlapping part of the two regions, A and B, by the set part of the two regions and the global information of interest. In this paper, A is the original image and B is the predicted image.

$$L_{\text{IoU}}^{\omega}(A, B) = 1 - \frac{\sum_{i=1}^H \sum_{j=1}^W (gt_{ij} * p_{ij}) * \omega_{ij}}{\sum_{i=1}^H \sum_{j=1}^W (gt_{ij} + p_{ij} - gt_{ij} * p_{ij}) * \omega_{ij}} \quad (6)$$

where L_{BCE}^{ω} represents the weighted BCE [15], the cross-entropy calculation function in the two-class classification task, and the pixel-level loss function.

$$L_{\text{BCE}}^{\omega} = \frac{\sum_{i=1}^H \sum_{j=1}^W \omega_{ij} (gt_{ij} * \log p_{ij} + (1 - gt_{ij}) * \log(1 - p_{ij}))}{\sum_{i=1}^H \sum_{j=1}^W \omega_{ij}} \quad (7)$$

where

$$\omega_{ij} = 1 + \gamma \alpha_{ij}$$

$$\alpha_{ij} = \left| \frac{\sum_{m,n \in A_{ij}^1} gt_{mn}}{\sum_{m,n \in A_{ij}^1} 1} - gt_{ij} \right| + \left| \frac{\sum_{m,n \in A_{ij}^2} gt_{mn}}{\sum_{m,n \in A_{ij}^2} 1} - gt_{ij} \right| + \left| \frac{\sum_{m,n \in A_{ij}} gt_{mn}}{\sum_{m,n \in A_{ij}} 1} - gt_{ij} \right|$$

where $gt_{ij} \in [0, 1]$ represents the original image: if the pixel (i, j) is the GBM, $gt_{ij} = 1$; otherwise $gt_{ij} = 0$; $p_{ij} \in [0, 1]$ represents the probability that pixel (i, j) is the GBM; ω_{ij} represents the weight value; α_{ij} indicates the importance of pixels by calculating the difference between pixels and their surrounding pixels; A_{ij} , A_{ij}^1 , and A_{ij}^2 represent the neighborhood of the pixel (i, j); and we used A_{ij} as the 31×31 rectangle, A_{ij}^1 as the 31×1 rectangle, and A_{ij}^2 as the 1×31 rectangle. By using these rectangles, we can consider and locate the GBM like a doctor would by taking the tissues around the GBM as the reference object; γ is a super parameter, and the value is 5 in this paper.

To enhance the segmentation of the GBM, the SSPNet integrates strip pooling to capture the striped features of the GBM and a reverse attention mechanism to highlight edge information. Strip pooling, with window sizes of (H, 1) or (1, W), averages features along rows or columns, effectively capturing long-range dependencies. The reverse attention mechanism dynamically enhances boundary details, thereby improving segmentation accuracy. In the experimental section, work was carried out to boost model transparency

and credibility. SSPNet employs feature map visualization to intuitively display how the model processes and extracts features. Quantitative analysis using metrics such as the Dice coefficient and IoU provided a clear understanding of the model's segmentation accuracy. Comparative experiments with other advanced medical image segmentation methods demonstrated SSPNet's superiority in segmenting the GBM. Integration with physical implications through quantitative data extraction, such as calculating the GBM thickness using skeleton extraction, enhanced the model's practicality and decision-making credibility. These methods have ensured SSPNet's effectiveness and interpretability in practical applications.

4. Experimental Design and Analysis

4.1. Experimental Details

GBM Datasets and Baselines: The experimental data used in this paper came from 700 electron microscopic images of renal biopsies, each with a resolution of 2048×2048 pixels, procured from 347 patients at the Southern Medical University. The dataset encompassed a diverse range of renal pathologies, including, but not limited to, primary chronic glomerular nephropathy (IgA nephropathy), minimal change disease (MCD), membranous nephropathy (MN), thin basement membrane nephropathy, diabetic nephropathy, and light mesangial proliferative glomerulonephritis.

To ensure the comprehensiveness and balance of the dataset, we analyzed the distribution of disease types and pathological variations across the 347 patients. The dataset included the following disease types and their respective sample counts:

- Primary chronic glomerular nephropathy (IgA Nephropathy): 74 patients (21.3%);
- Minimal change disease (MCD): 54 patients (15.6%);
- Membranous nephropathy (MN): 59 patients (17%);
- Thin basement membrane nephropathy: 26 patients (7.5%);
- Diabetic nephropathy: 33 patients (9.5%);
- Light mesangial proliferative glomerulonephritis: 53 patients (15.3%);
- Lupus nephritis: 48 patients (13.8%).

This distribution ensures that the dataset covered a wide spectrum of renal pathologies, minimizing potential bias toward any specific disease type. Additionally, the dataset included a variety of pathological variations within each disease category, such as different stages of glomerular basement membrane thickening, mesangial proliferation, and podocyte effacement. This diversity allowed the model to generalize well across different pathological conditions, and ensures robust performance in real-world clinical scenarios.

The original images were marked by the pathologist. Figure 1 shows an example of the dataset and annotation. The annotated information of the original data included compacts, foot processes, endothelial cells, mesangial area, and the GBM. This paper focused on extracting the thickness of the GBM, so only the annotated information regarding the GBM was selected. Within this information, there were no abnormal changes in the basement membrane of MCD, IgA, or mild mesangial proliferative glomerulonephritis.

Despite the vast amount of material available, there is little research on the semantic segmentation of the GBM. Therefore, in our research, we conducted a comparative analysis of our novel SSPNet approach against five state-of-the-art (SOTA) medical image segmentation methodologies, aiming to evaluate its performance and identify potential advantages. These other methodologies were Deeplabv3 [20], Deeplabv3+ [24], HarDNet-MSEG [23], F3Net [25], and PraNet [26].

Training Settings and Metrics: During the experimental phase, we randomly partitioned the dataset into three distinct subsets: a training set comprising 500 images, a validation set with 100 images, and a test set containing 100 images. Instead of utilizing data augmentation, we adopted a multiscale training approach with scales of 0.75, 1, and 1.25 to enhance model robustness.

For quantitative assessment, we adhered to the evaluation metrics employed in prior works [27,28], namely, the mean Dice coefficient and the mean Intersection over Union (IoU). To gain a more comprehensive understanding of our model's performance, we additionally incorporated three supplementary metrics: S-measure (Sm), which assesses structural similarity between predictions and ground truths at both the regional and object levels [29]; weighted F-measure (Fm), a holistic performance indicator that balances weighted precision and recall [30]; mean absolute error (MAE), which quantifies the average pixel-wise difference between predicted and ground truth saliency maps [31].

Implementation Details The model in this paper was implemented under the Pytorch framework. We used six GPUs, which were NVIDIA GeForce RTX 3090 or NVIDIA TITANXp, sourced from NVIDIA Corporation (Santa Clara, CA, USA). The experimental operating system used was ubuntu14.04. To refine the optimization process, we adopted the Adam optimizer, tuning the overall parameters with a learning rate set at 3×10^{-4} . The entire network underwent end-to-end training, utilizing a batch size of 2 across 60 epochs. All input data were consistently resized to 1024×1024 pixels to ensure uniformity. Subsequently, the performance of the proposed network architecture was benchmarked against alternative network configurations, facilitating a comprehensive comparative analysis.

To evaluate the real-time performance of our model, we measured the inference speed on a single NVIDIA GeForce RTX 3090 GPU. The average inference time per image was approximately 20 milliseconds (ms), which translates to a frame rate of approximately 50 frames per second (FPS). This high frame rate ensures that our model can process video streams in real time, making it suitable for practical applications such as automated pathological diagnosis in clinical settings.

4.2. Comparison of Pooling Operations

This paper compared the SSP model with the strip pooling module (SPM), the mixed pooling module (MPM), and the SPM + MPM models in SPNet. The experimental results are shown in Table 1. Our SSPNet surpassed the SPNet by a notable margin, achieving a mean Dice score improvement of over 1%. This observation underscores the model's enhanced capacity for accurately segmenting intricate strip features, indicative of its superior learning proficiency. Therefore, better experimental results were obtained.

Table 1. Comparison of pooling operations.

Model	Dice↑	IoU↑	Sm↑	Fm↑	MAE↓
Original (Res2Net)	0.746	0.603	0.806	0.710	0.042
Res2Net + SPM	0.748	0.604	0.813	0.709	0.043
Res2Net + MPM	0.747	0.603	0.810	0.699	0.044
Res2Net + SPM + MPM	0.738	0.594	0.811	0.687	0.045
Res2Net + SSP	0.756	0.615	0.818	0.709	0.042

Note: ↑ indicates higher values are better, ↓ indicates lower values are better.

In Figure 4, we provide the GBM segmentation results of SSPNet. Green indicates professional labeling of nephropathy, red indicates automatic labeling by the algorithm, and yellow indicates overlapping parts. It can be seen from Figure 4 that the strip information marked by SSP, especially the part framed by the blue box, is more comprehensive, and the difference is obvious.

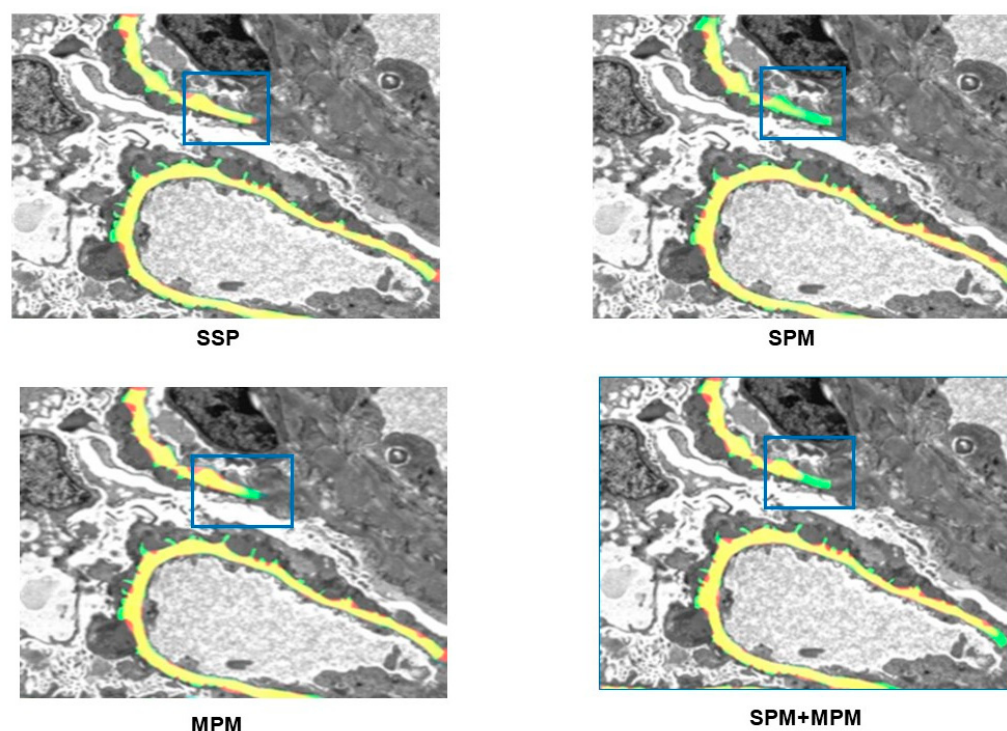


Figure 4. Qualitative results of the different methods. Green indicates professional labeling of nephropathy, red indicates automatic labeling by the algorithm, and yellow indicates overlapping parts. The blue boxes visually emphasize the notable differences in the outcomes produced by different methodologies.

4.3. Ablation Experiment

To verify the feasibility of each step in this method, ablation experiments were carried out in the specific optimization stage of this method to verify the correctness and necessity of each processing stage.

We first tested each module in isolation to verify its influence. As shown in Table 2, Res2Net was selected as the backbone of this paper. Res2Net + Dilated means Dilated was added to Res2Net. Res2Net + SSP refers to an SSP module being added to extract mixed textures for the basic network framework. Res2Net + AT indicates that an attention module was added to the basic network to enhance boundary information. Res2Net + LossA indicates that self-adaptive loss was used to replace the original loss. Additionally, we tested the overlapping influence by gradually adding the module. SSPNet is the proposed method.

Table 2. Ablation study for SSPNet on the GBM dataset.

Model	Dice↑	IoU↑	Sm↑	Fm↑	MAE↓
Res2Net	0.746	0.603	0.806	0.710	0.042
+Dilated	0.773	0.637	0.812	0.742	0.038
+SSP	0.756	0.615	0.818	0.709	0.042
+AT	0.767	0.630	0.813	0.737	0.039
+LossA	0.772	0.637	0.818	0.716	0.041
+Dilated + SSP	0.776	0.642	0.821	0.730	0.039
+Dilated + SSP + AT	0.785	0.653	0.828	0.741	0.037
SSPNet	0.792	0.662	0.836	0.753	0.036

It can be seen in Table 2 that the incorporation of dilated convolutions, facilitating exponential enlargement of the receptive field while maintaining resolution and coverage,

resulted in an accuracy of 0.773, demonstrating an effective approach for the enhancement of performance. At the same time, the addition of an attention mechanism that emphasized boundary information also improved the performance of the final network. The accuracy was increased to 0.767. Then, in view of the strip texture features of the GBM, an SSP module was directly added on the basis of Res2Net to emphasize the extraction of strip texture features, and an accuracy rate of 0.756 was obtained. This shows that the SSP module has a good effect when processing images with striping problems. We changed the loss function and obtained an accuracy of 0.772. Finally, we gradually integrated the module, obtaining an accuracy of 0.792.

4.4. Performance Comparison

It can be found in Table 3 that for strip images, such as those of the GBM, the SSP module can extract the strip texture, and the effect is better than that of other methods. Figure 5 shows the specific segmentation results. It can be intuitively observed that this method can predict the strip texture well. Finally, automatic thickness measurement of the GBM is helpful for pathologists in carrying out subsequent pathological diagnoses. We used the close morphological operation to remove some noise and then used the skeleton algorithm [29,30] to extract the skeleton of the GT graph. Taking the longest line segment in the skeleton as the central axis of the area, the distance from the point on the central axis to the background pixel was calculated and the thickness of the GBM obtained, providing quantitative information for doctors' later diagnoses. We obtained digital features and drew a digital feature graph, which is shown in Figure 5.

Table 3. Comparison of the results from the various models.

Model	Dice $\uparrow$	IoU $\uparrow$	Sm $\uparrow$	Fm $\uparrow$	MAE $\downarrow$
Deeplabv3 [20]	0.760	0.625	0.819	0.730	0.040
Deeplabv3+ [24]	0.754	0.621	0.814	0.711	0.041
HarDNet-MSEG [23]	0.763	0.629	0.823	0.704	0.042
F3Net [25]	0.781	0.652	0.831	0.731	0.038
PraNet [26]	0.785	0.656	0.826	0.754	0.037
Our method	0.792	0.662	0.836	0.753	0.036

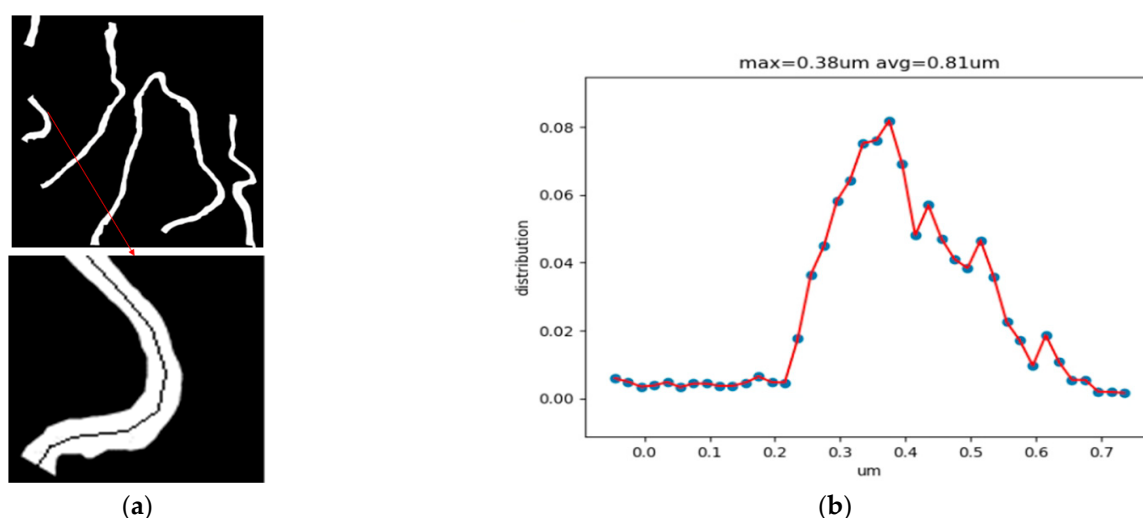


Figure 5. The thickness of the GBM was calculated using its skeleton GBM. (a) The skeleton of the GBM; (b) the thickness of the GBM. Red lines highlight magnified regions of key details, and blue dots mark the thickness measurement points of GBM.

To assess the statistical significance of the performance differences between SSPNet and other state-of-the-art models, we conducted paired *t*-tests on the Dice coefficients obtained from the test set. *p*-values and 95% confidence intervals (CIs) were calculated to quantify the significance of the differences (Table 4).

Table 4. Statistical significance analysis of the Dice coefficients.

Model	Mean Dice Coefficient	<i>p</i> -Value vs. SSPNet	95% CI of the Difference
Deeplabv3	0.760	<0.001	[0.012, 0.036]
Deeplabv3+	0.754	<0.001	[0.018, 0.042]
HarDNet-MSEG	0.763	<0.001	[0.015, 0.039]
F3Net	0.781	0.012	[0.003, 0.021]
PraNet	0.785	0.008	[0.002, 0.018]
SSPNet	0.792	-	-

Note: *p*-values indicate the significance of the differences between SSPNet and the other models. A *p*-value < 0.05 suggests a statistically significant difference. The 95% CIs of the differences show the range within which the true difference in the mean Dice coefficients is likely to fall.

The statistical analysis revealed that SSPNet achieved significantly higher Dice coefficients compared to Deeplabv3, Deeplabv3+, HarDNet-MSEG, F3Net, and PraNet (*p*-values < 0.001 for Deeplabv3, Deeplabv3+, and HarDNet-MSEG; *p*-value = 0.012 for F3Net; *p*-value = 0.008 for PraNet). The 95% confidence intervals of the differences further support the robustness of SSPNet's performance, indicating that the improvements are not due to random variation.

4.5. Experiments on Polyp Segmentation

In conclusion, the proposed model's versatility was validated by testing on a publicly available polyp dataset, demonstrating its applicability beyond GBM detection tasks. To uphold a rigorous verification process, we adhered strictly to the training protocol stipulated in [26], where the Kvasir and CVC-ClinicDB (also known as CVC-612) image datasets were systematically partitioned, with 80% used for training, 10% for validation, and the remaining 10% for testing, and with all assignments being carried out randomly to ensure unbiased evaluation.

As evidenced by the results presented in Tables 5 and 6, our methodology yielded favorable outcomes when applied to the polyp datasets, underscoring its robust performance and broad applicability in medical image semantic segmentation. Notably, our model, originally conceived for addressing strip structures, demonstrated its capability to deliver satisfactory results on generic datasets as well. This observation leads us to speculate that the employment of strip pooling in our approach facilitates the capture of long-range dependencies within the data, contributing to its enhanced performance and versatility.

Table 5. Comparison with the results of various models on Kvasir [28], with U-Net, U-Net++, ResU-Net, ResU-Net, and PraNet evaluation scores from [10]. Nan denotes that the results are unavailable.

Model	Dice↑	IoU↑	Sm↑	Fm↑	MAE↓
U-Net [20]	0.818	0.444	0.858	0.794	0.055
U-Net++ [24]	0.821	0.743	0.862	0.808	0.048
ResU-Net [23]	0.791	Nan	Nan	Nan	Nan
ResU-Net++ [25]	0.813	0.793	Nan	Nan	Nan
PraNet [26]	0.898	0.840	0.915	0.885	0.030
Ours (SASPNet)	0.908	0.859	0.924	0.909	0.018

Table 6. Comparison with the results of various models on CVC-612 [32], with U-Net, U-Net++, ResU-Net, ResU-Net, and PraNet evaluation scores from [10]. Nan denotes that the results are unavailable.

Model	Dice↑	IoU↑	Sm↑	Fm↑	MAE↓
U-Net [20]	0.823	0.755	0.954	0.889	0.019
U-Net++ [24]	0.794	0.729	0.931	0.873	0.022
ResU-Net [23]	0.779	Nan	Nan	Nan	Nan
ResU-Net++ [25]	0.796	0.796	Nan	Nan	Nan
PraNet [26]	0.899	0.849	0.979	0.896	0.009
Ours (SASPNet)	0.915	0.866	0.933	0.914	0.011

4.6. Dataset Specificity and Generalizability

The dataset used in this study comprised 700 electron microscopic images of renal biopsies from 347 patients, covering a diverse range of renal pathologies, including IgA nephropathy, minimal change disease, and membranous nephropathy. The dataset's diversity in terms of pathology types and variations ensured that SSPNet was trained and tested on a wide range of scenarios, which is crucial for robust performance. Additionally, the dataset's balanced distribution of pathologies ensured that SSPNet learned generalized features rather than being overfit to specific disease types.

However, the images exhibited the following unique characteristics: high resolution (2048×2048 pixels) and low contrast, specifically between the GBM and surrounding tissues. Such attributes inherently influenced SSPNet's performance. These characteristics may also have introduced limitations: models trained on high-resolution EM images might struggle with lower-resolution datasets, and dependency on edge attention could reduce effectiveness in high-contrast scenarios. While SSPNet demonstrated high precision in segmenting the GBM in this dataset, it is essential to evaluate its generalizability to other datasets. The observed improvements in Dice coefficients and IoU metrics were significant within the context of this study, but they may not be directly transferable to datasets with different imaging modalities, resolutions, or pathological conditions. For instance, datasets with higher contrast or different tissue structures might yield different performance metrics.

To evaluate generalizability, we tested SSPNet on the Kvasir and CVC-612 polyp segmentation datasets, which differ significantly in imaging modality, resolution, and target structures. As shown in Tables 5 and 6, SSPNet achieved competitive Dice scores (0.908 and 0.915, respectively), outperforming established methods like PraNet. These results suggest that SSPNet's design—particularly its strip pooling mechanism for capturing long-range dependencies and its adaptive loss function for emphasizing challenging regions—enhances its adaptability to diverse medical imaging tasks. However, the observed gains in GBM segmentation (Dice = 0.792) are partially context specific. For instance, the strip pooling module excels in segmenting elongated structures (e.g., the GBM), but may offer only marginal benefits for compact or irregularly shaped targets.

5. Conclusions

Accurate semantic segmentation of the glomerular basement membrane (GBM) and automatic measurements of its thickness are crucial for aiding pathologists in conducting follow-up pathological diagnoses. In this study, we proposed the self-adaptive strip pooling network (SSPNet) to address the challenges posed by the complex ultrastructure and irregular shape of the GBM. By leveraging the strip-like features of the GBM, SSPNet effectively extracted both strip and square textures through its innovative refined strip pooling (RSP) module. Additionally, the incorporation of a reverse attention mechanism significantly enhanced edge extraction, thereby improving segmentation accuracy. Through extensive ex-

periments and comparisons with state-of-the-art methods, SSPNet demonstrated superior performance in GBM segmentation, achieving high precision and robustness.

However, despite its notable achievements, SSPNet has certain limitations. As shown in Figure 6, segmentation errors still occurred in some cases. One possible reason is the reliance solely on textural information without incorporating expert knowledge, which may be insufficient for accurate classification in all scenarios. Moreover, the manual annotations used for training and evaluation, performed by expert pathologists, are subject to inter-observer variability, a common issue in medical image analysis. This variability could subtly influence model training and evaluation results. Future studies should conduct inter-rater reliability analyses, such as using Cohen's kappa or Dice agreement scores, across multiple annotators to quantify this effect. Furthermore, real-world clinical practices may involve differing diagnostic criteria or annotation granularity. To improve reproducibility, we recommend adopting consensus-driven annotation guidelines and integrating semi-automated tools to reduce human subjectivity [33]. These measures would strengthen SSPNet's reliability and facilitate its application in diverse clinical workflows.

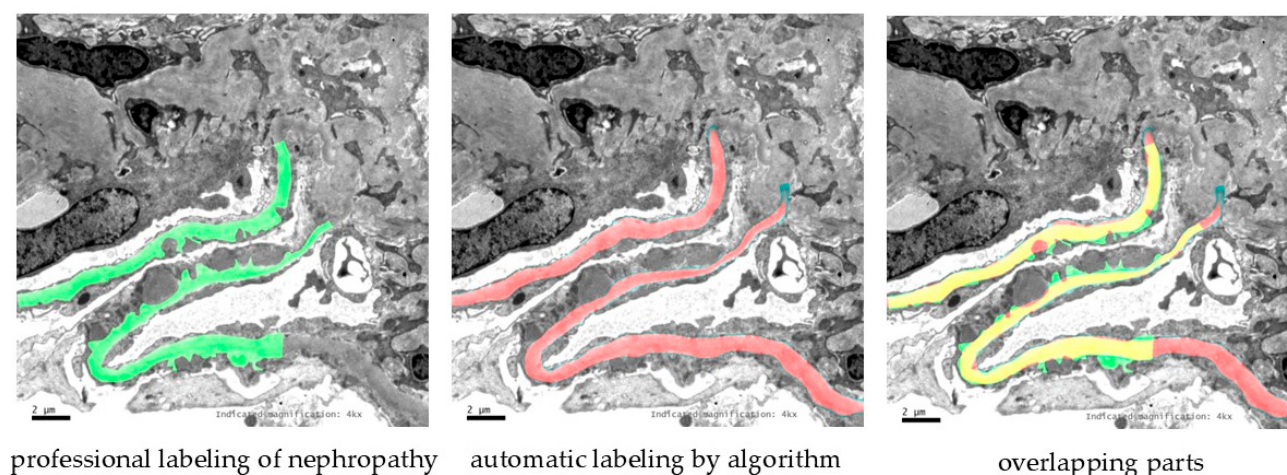


Figure 6. Green indicates professional labeling of nephropathy, red indicates automatic labeling by the algorithm, and yellow indicates overlapping parts.

The proposed method's performance on external polyp datasets (Tables 5 and 6) suggests promising generalizability. However, its efficacy in real-world GBM segmentation depends on addressing the aforementioned challenges. For broader adoption, validation across multi-institutional cohorts and imaging modalities (e.g., optical microscopy or other EM protocols) is imperative. Collaborative efforts with clinical partners will be essential to refine SSPNet's adaptability to practical diagnostic scenarios while ensuring ethical and regulatory compliance.

In the future, the following studies will be conducted:

- (1) **Automated Data Annotation:** Investigation of methods for automated data annotation to improve the accuracy and efficiency of the training process. By developing algorithms that can automatically generate high-quality annotations, we aim to reduce the reliance on manual annotation and enhance the scalability of our approach.
- (2) **Multi-class Segmentation and Expert Knowledge Integration:** Extend the segmentation framework to support multi-class segmentation tasks, enabling the simultaneous identification of multiple structures within renal biopsy images. Additionally, we will explore ways to incorporate expert knowledge and clinical guidelines into the segmentation process. By integrating expert experience and domain-specific rules, we

aim to improve the accuracy and reliability of the segmentation results, making them more aligned with clinical needs.

Author Contributions: Conceptualization, C.S. and X.H.; methodology, C.S.; software, X.L.; validation, C.S. and X.L.; writing—original draft preparation, C.S.; writing—review and editing, X.H.; visualization, X.L.; project administration, C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Shanxi Province, China, grant number 20210302124257.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available only upon request from the corresponding author, due to the fact that they are sourced from a third-party provider, whose terms of use and data-sharing policies strictly prohibit the unauthorized disclosure or public dissemination of the information. As the researcher utilizing these data, the author is bound by contractual obligations and ethical guidelines to respect the privacy and confidentiality agreements established by the data owner. Therefore, direct access to the raw data is limited to ensure compliance with these regulations and to maintain the integrity of the data-sharing framework. However, for purposes of verification, collaboration, or further analysis, the corresponding author is willing to facilitate requests for data access, subject to the approval of the third-party data provider and adherence to all relevant legal and ethical requirements.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Levey, A.S.; Coresh, J.; Bolton, K.; Culleton, B.; Harvey, K.S.; Ikizler, T.A.; Johnson, C.A.; Kausz, A.; Kimmel, P.L.; Kusek, J.; et al. K/DOQI Clinical Practice Guide lines for Chronic Kidney Disease: Evaluation, Classification and Stratification National Kidney Foundation. *Am. J. Kidney Dis.* **2003**, *139*, 137–147.
2. Bondugulapati, L.N.R.; Shandilya, S. Chronic kidney disease and cardiovascular disease. *Curr. Opin. Lipidol.* **2015**, *26*, 353–354. [[CrossRef](#)] [[PubMed](#)]
3. Ong, S.H.; Giam, S.T.; Jayasooriah; Sinniah, R. Adaptive window-based tracking for the detection of membrane structures in kidney electron micrographs. *Mach. Vis. Appl.* **1993**, *6*, 215–223. [[CrossRef](#)]
4. Kamenetsky, I.; Rangayyan, R.M.; Benediktsson, H. Analysis of the glomerular basement membrane in images of renal biopsies using the split-and-merge method: A pilot study. *J. Digit. Imaging* **2010**, *23*, 463–474. [[CrossRef](#)] [[PubMed](#)]
5. Rangayyan, R.M.; Kamenetsky, I.; Benediktsson, H. Segmentation and analysis of the glomerular basement membrane in renal biopsy samples using active contours: A pilot study. *J. Digit. Imaging* **2010**, *23*, 323–331. [[CrossRef](#)] [[PubMed](#)]
6. Wu, H.S.; Dikman, S.; Gil, J. A semi-automatic algorithm for measurement of basement membrane thickness in kidneys in electron microscopy images. *Comput. Methods Programs Biomed.* **2010**, *97*, 223–231. [[CrossRef](#)] [[PubMed](#)]
7. Wu, H.S.; Dikman, S. Segmentation and thickness measurement of glomerular basement membranes from electron microscopy images. *J. Electron Microsc.* **2010**, *59*, 409–418. [[CrossRef](#)] [[PubMed](#)]
8. Liu, J. Implementation of A Semi-Automatic Tool for Analysis of TEM Images of Kidney Samples. Master's Thesis, Uppsala University, Uppsala, Sweden, 2012.
9. Li, C.; Lu, Y.; Li, M.; Li, M.; Li, R.; Cao, L. Automatic segmentation of glomerular basement membrane based on image patch matching. *J. Comput. Appl.* **2016**, *36*, 3201.
10. Mu, L.I.; Chuang-Quan, L.I.; Lei, C.; Ying-Ying, C.; Yuan-Yuan, L.; Yan-Meng, L.U. Basement membrane segmentation of glomerular tem images based on particle filter. *Sci. Technol. Andengineering* **2017**, *17*, 30–36.
11. Guo, Y.; Ashour, A.S.; Sun, B. A novel glomerular basement membrane segmentation using neutrosophic set and shearlet transform on microscopic images. *Health Inf. Sci. Syst.* **2017**, *5*, 15. [[CrossRef](#)] [[PubMed](#)]
12. Cao, L.; Lu, Y.; Li, C.; Yang, W. Automatic segmentation of pathological glomerular basement membrane in transmission electron microscopy images with random forest stacks. *Comput. Math. Methods Med.* **2019**, *2019*, 1684218. [[CrossRef](#)] [[PubMed](#)]
13. Taghanaki, S.A.; Abhishek, K.; Cohen, J.P.; Cohen-Adad, J.; Hamarneh, G. Deep semantic segmentation of natural and medical images: A review. *Artif. Intell. Rev.* **2020**, *54*, 137–178. [[CrossRef](#)]

14. Alsaleh, A.M.; Albalawi, E.; Algosaihi, A.; Albakheet, S.S.; Khan, S.B. Few-Shot Learning for Medical Image Segmentation Using 3D U-Net and Model-Agnostic Meta-Learning (MAML). *Diagnostics* **2024**, *14*, 1213. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Hou, Q.; Zhang, L.; Cheng, M.-M.; Feng, J. Strip Pooling: Rethinking Spatial Pooling for Scene Parsing. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 14–19 June 2020; pp. 4002–4011.
16. Gao, S.-H.; Cheng, M.-M.; Zhao, K.; Zhang, X.-Y.; Yang, M.-H.; Torr, P.H. Res2Net: A new multi-scale backbone architecture. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 652–662. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Sandler, M.; Baccash, J.; Zhmoginov, A.; Howard, A. Non-discriminative data or weak model? On the relative importance of data and model resolution. *arXiv* **2019**, arXiv:1909.03205.
18. Huang, Z.; Wang, X.; Wei, Y.; Huang, L.; Shi, H.; Liu, W.; Huang, T.S. CCNet: Criss-cross attention for semantic segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019. [\[CrossRef\]](#)
19. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv* **2017**, arXiv:1706.05587. [\[CrossRef\]](#)
21. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.
22. Lazebnik, S.; Schmid, C.; Ponce, J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), New York, NY, USA, 17–22 June 2006; Volume 2, pp. 2169–2178.
23. Huang, C.H.; Wu, H.Y.; Lin, Y.L. HarDNet-MSEG: A Simple Encoder-Decoder Polyp Segmentation Neural Network that Achieves over 0.9 Mean Dice and 86 FPS. *arXiv* **2021**, arXiv:2101.07172.
24. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. *Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation*; Springer: Cham, Switzerland, 2018.
25. Wei, J.; Wang, S.; Huang, Q. F3Net: Fusion, Feedback and Focus for Salient Object Detection. In Proceedings of the AAAI Conference on Artificial Intelligence 2019, Honolulu, HI, USA, 27 January–1 February 2019; Volume 34, pp. 12321–12328.
26. Fan, D.P.; Ji, G.P.; Zhou, T.; Chen, G.; Fu, H.; Shen, J.; Shao, L. PraNet: Parallel Reverse Attention Network for Polyp Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, 4–8 October 2020; Springer International Publishing: Cham, Switzerland, 2020; pp. 263–273.
27. Jha, D.; Smedsrud, P.H.; Riegler, M.A.; Johansen, D.; De Lange, T.; Halvorsen, P.; Johansen, H.D. Resunet++: An advanced architecture for medical image segmentation. In Proceedings of the 2019 IEEE International Symposium on Multimedia (ISM), San Diego, CA, USA, 9–11 December 2019; pp. 225–2255.
28. Fan, D.-P.; Cheng, M.-M.; Liu, Y.; Li, T.; Borji, A. Structure-Measure: A New Way to Evaluate Foreground Maps. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 4558–4567.
29. Margolin, R.; Zelnik-Manor, L.; Tal, A. How to evaluate foreground maps? In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 248–255.
30. Perazzi, F.; Krähenbühl, P.; Pritch, Y.; Hornung, A. Saliency filters: Contrast based filtering for salient region detection. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, USA, 16–21 June 2012; pp. 733–740.
31. Saeed, K.; Tabedzki, M.; Rybnik, M.; Adamski, M. K3M: A universal algorithm for image skeletonization and a review of thinning techniques. *Int. J. Appl. Math. Comput. Sci.* **2010**, *20*, 317–335. [\[CrossRef\]](#)
32. Bernal, J.; Sánchez, F.J.; Fernández-Esparrach, G.; Gil, D.; Rodríguez, C.; Vilariño, F. Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Comput. Med. Imaging Graph.* **2015**, *43*, 99–111. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Liu, X.; Shi, G.; Wang, R.; Lai, Y.; Zhang, J.; Sun, L.; Yang, Q.; Wu, Y.; Li, M.; Han, W.; et al. Feature-Prompting GBMSeg: One-Shot Reference Guided Training-Free Prompt Engineering for Glomerular Basement Membrane Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Marrakesh, Morocco, 6–10 October 2024; Springer: Cham, Switzerland, 2024. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.