Computerized Medical Imaging and Graphics 100 (2022) 102104

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

Boundary-aware glomerulus segmentation: Toward one-to-many stain generalization

Jefferson Silva [a,b], Luiz Souza [b], Paulo Chagas [b], Rodrigo Calumby [d], Bianca Souza [b,c], Izabelle Pontes [b], Angelo Duarte [d], Nathanael Pinheiro [e], Washington Santos [b,c], Luciano Oliveira [b,*]

[a] Universidade Federal do Maranhão, Brazil
[b] Universidade Federal da Bahia, Brazil
[c] Fundação Oswaldo Cruz, Brazil
[d] Universidade Estadual de Feira de Santana, Brazil
[e] Imagepat Laboratory, Brazil

ARTICLE INFO

Keywords:
Kidney
Segmentation
Molecular
Cellular imaging
End-to-end learning in medical imaging

ABSTRACT

The growing availability of scanned whole-slide images (WSIs) has allowed nephropathology to open new possibilities for medical decision-making over high-resolution images. Diagnosis of renal WSIs includes locating and identifying specific structures in the tissue. Considering the glomerulus as one of the first structures analyzed by pathologists, we propose here a novel convolutional neural network for glomerulus segmentation. Our end-to-end network, named DS-FNet, combines the strengths of semantic segmentation and semantic boundary detection networks via an attention-aware mechanism. Although we trained the proposed network on periodic acid-Schiff (PAS)-stained WSIs, we found that our network was capable to segment glomeruli on WSIs stained with different techniques, such as periodic acid-methenamine silver (PAMS), hematoxylin-eosin (HE), and Masson trichrome (TRI). To assess the performance of the proposed method, we used three public data sets: HuBMAP (available in a Kaggle competition), a subset of the NEPTUNE data set, and a novel challenging data set, called WSI_Fiocruz. We compared the DS-FNet with six other deep learning networks: original U-Net, our attention version of U-Net called AU-Net, U-Net++, U-Net3Plus, ResU-Net, and DeepLabV3+. Results showed that DS-FNet achieved equivalent or superior results on all data sets: On the HuBMAP data set, it reached a dice score (DSC) of 95.05%, very close to the first place (95.15%); on the NEPTUNE and WSI_Fiocruz data sets, DS-FNet obtained the highest average DSC, whether on PAS-stained images or images stained with other techniques. To the best we know, this is the first work to show consistently high performance in a one-to-many-stain glomerulus segmentation following a thorough protocol on data sets from different medical labs.

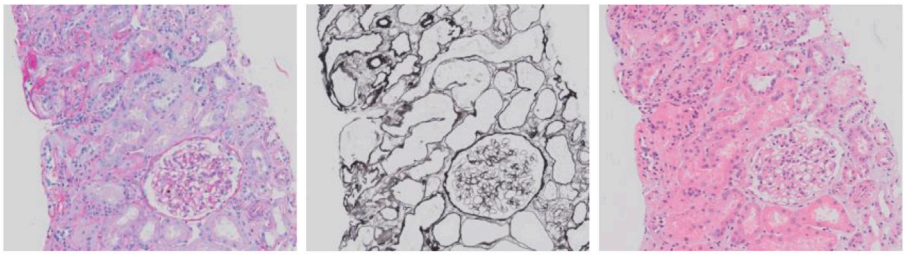
1. Introduction

Renal biopsy represents the most common and standard way to diagnose several kidney diseases (Bandari et al., 2016). The conventional approach to prepare a renal biopsy is to employ stained tissue sections on histological slides. The slides can be examined under the microscope or scanned, generating high-resolution whole slide images (WSIs). The use of digital images to analyze samples of tissues allowed booming the advances in the field of digital pathology whose gains come with the more standard and easy-to-share material collection and processing (Barisoni et al., 2017). This growing availability of WSIs facilitated the collaboration among pathologists and the composition of digital libraries. While the job of localization and identification of renal structures is the essential in the routine of pathologists, it is expected that diagnosing over digital images will promote breakthroughs in the field, especially through the consolidation of the computational nephropathology as a new research field.

Among the structures present in the human kidney, the glomerulus is the one responsible for blood filtration (Weinstein and Anderson, 2010). Due to this filtering function, primary or systemic diseases causing lesions in the glomerulus can lead to renal failure with high burden for the patient and public health systems. Because of the impact in the diagnosis pipeline, glomeruli are usually one of the first structures to be assessed by pathologists (Bellur et al., 2019).

* Corresponding author.
E-mail address: lrebouca@ufba.br (L. Oliveira).

https://doi.org/10.1016/j.compmedimag.2022.102104
Received 14 February 2022; Received in revised form 28 May 2022; Accepted 8 July 2022
Available online 12 August 2022



(a) PAS (b) PAMS (c) HE

Fig. 1. Image samples from WSI_Fiocruz data set in different stains: Periodic acid-Schiff (PAS), periodic acid-methenamine silver (PAMS), and hematoxylin-eosin (HE).

Quantitative and qualitative evaluations of these structures are complex and time-consuming tasks, reaching sometimes poor agreement among pathologists. In this context, automatic segmentation pipelines can help pathologists streamline the identification of relevant regions on WSIs and potentially leverage rapid kidney diagnoses and prognosis prediction.

For glomerulus localization and segmentation, due to the inherent issues involving image annotation (*e.g.*, time-consuming, specialist-dependent), pre-clinical works use machine-learning (ML) approaches considering only one stain for training and evaluation have been used (see Table 1). By building ML experiments for one-to-one-stain evaluation, high effectiveness may indeed be easier to achieve because training and validation sets have the same data distribution. Training ML models in a one-to-one multi-stain scenario (one model for each stain) can raise some limitations. Constraints such as time, specialists, and computational resources are likely to occur in several nephropathology laboratories, which might bring difficulties in gathering a vast data collection in sets of different stains. Without a reasonable amount of images for every staining technique, the underrepresented ones might result in a poorly trained model, which ultimately lacks robustness in a one-to-many stain evaluation. In such context, a question naturally arises: *what if we could use a large amount of images stained by a unique technique to learn features for other stained images?* To answer this question, we started with the analysis of the same glomerulus viewed in different stains, as illustrated in Fig. 1. Although differently stained, the glomerulus seems to preserve the same characteristics on the edges, which delineate either the Bowman's capsule or its internal structures (refer to Section 3 for more details on this medical topic).

Considering a convolutional neural network (CNN) architecture, leveraged by attention techniques employed in strategical places in the CNN to explore the boundary information, we introduce here a novel deep-learning (DL) architecture, called DS-FNet, that achieved promising results when evaluated on data sets gathered in different labs – thus captured with different instrumentation – while stained in different ways. Experiments were carried out on three data sets: HubMAP (Borner et al., 2021), a subset of the NEPTUNE data set (Jayapandian et al., 2021), and our novel data set called WSI_Fiocruz. All data sets together contain 660 WSIs and 5309 glomeruli. Our DL architecture was trained on the HubMAP training data set, containing 3568 periodic acid-Schiff (PAS)-stained WSIs, due to the lack of a substantial number of images for training on other stained WSIs. Our experiments showed that DS-FNet reached equivalent or superior results in the generalization of glomerulus segmentation, considering a one-to-many-stain segmentation, when compared to six other DL networks: original U-Net, AU-Net (our attention version of U-Net), U-Net++ (Zhou et al., 2018), U-Net3Plus (Huang et al., 2020), ResU-Net (Diakogiannis et al., 2020), and DeepLabV3+ (Diakogiannis et al., 2020).

2. State-of-the-art and contributions

Although there are several works that cope with the problem of WSI components segmentation, none of them solves the problem of one-to-many-stain glomerulus segmentation. To tackle this problem, we propose here a novel fusion-based, end-to-end DL architecture and a novel one-to-many-stain data set. We highlight these two novelty aspects as follows, bringing our contributions regarding the related works.

In the pathology practice, different stains can be applied to tissue samples in such a way that each stain stands out different morphological components of the renal structures. The choice of the stain depends on the problem to be assessed, personal preference of pathologists, or the laboratory routine (Alturkistani et al., 2016; Bancroft and Gamble, 2008), directly influencing the segmentation results (see Section 3 for more details on this topic). Considering the morphological differences highlighted by each stain, a common ML approach is to train a neural network model using images of a specific stain, later performing the segmentation on images where that same stain was applied (one-to-one-stain segmentation). This type of approach is found in the majority of the related works summarized in Table 1, even for cases where more than one stain was considered. Another exception is the work of Jiang et al. (2021), which uses a many-to-many segmentation, mixing different stains in the sets of training and testing. We introduce here a novel one-to-many stain evaluation for glomerulus segmentation: instead of using images with the same stain for training and prediction (just as all the related works in Table 1), our proposed method shows that it is possible to obtain competitive results for multi-stain glomerular segmentation by training with only one stain. Next, we discuss the works summarized in Table 1 according to the segmentation approaches and data sets used for performance evaluation.

Segmentation approaches: All works adopt one or more DL architectures. DL became very popular for medical imaging tasks, also achieving state-of-the-art results in several medical domains (Seeja and Suresh, 2019; Guo et al., 2019; Sornapudi et al., 2020; Jayapandian et al., 2021). The majority of works uses the U-Net architecture or its variations for image segmentation. As these works achieved superior results in the segmentation of renal structures, we also adopted U-Net as our baseline architecture. In fact, 10 out of 18 works in Table 1 use only U-Net (Jayapandian et al., 2021; Davis et al., 2021; Hermsen et al., 2019; Jha et al., 2021; Bueno et al., 2020), variations of U-Net (Gadermayr et al., 2019; Bouteldja et al., 2021) or combine it with other methods (Mei et al., 2020; Zeng et al., 2020; de Bel et al., 2018). The remaining works explore other DL-based segmentation approaches such as one DL network: Mask-RCNN (Jiang et al., 2021) and DeepLabV2 (Lutnick et al., 2019; Ginley et al., 2020); two separate DL networks: MaskRCNN and FastRCNN (Altini et al., 2020a), and SegNet and DeepLabV3+ (Altini et al., 2020b); three separate DL networks: Mask-RCNN, U-Net, and DeepLabV3 (Jha et al., 2021); a combination of two DL networks: SegNet and AlexNet (Bueno et al., 2020); and finally pipelines that combine DL approaches with conventional image processing methods (Marsh et al., 2018; Kannan et al., 2019; Ginley et al., 2019).

A noteworthy work, not listed in Table 1, is the one found in Gallego et al. (2021), which addresses the problem of one-to-many-stain classification. That work relies upon a standard U-Net-based workflow

to detect and classify glomeruli in a WSI. There is no report on the segmentation performance, but the classification results align with our initial hypothesis for the one-to-many-stain segmentation.

Different from the previous work, we conceived an architecture based on a fusion of branches that are optimized all together to enrich the features of the segmentation network while being trained in an end-to-end fashion; all that is evaluated in a one-to-many-stain segmentation task.

Resources for segmentation assessment: Table 1 shows that our work considers the largest number of WSIs from different medical labs, when it evaluates glomerulus segmentation, and the fourth largest number of glomeruli. Due to the limited availability of public data sets and the time-consuming in the process of annotation by pathologists, most works build a data set comprising less than 100 WSIs and less than 3000 glomeruli.

Instead, we opted to use as a training set the data set that contains the stain with the greatest number of annotated glomeruli, subsequently testing our model in data sets from different medical labs.

Contributions: Even with differences in color, texture, and highlighted structures, the general shape of the glomerulus is mostly preserved across different stains. Our work seeks to develop a novel deep network based on a parallel branch to assist a baseline segmentation network on the task of segmenting the glomeruli found on WSIs. We hypothesize that by detecting boundaries, it is possible to capture better structural information than the features (mostly derived from color and texture) captured by regular segmentation networks. Hopefully, by bringing boundary information along with the features learned from the segmentation network, we can achieve generalization of the glomerular structure across different stains by training with just one stain. To allow the simultaneous learning of both segmentation and boundary detection tasks, we also adopted a set of attention mechanisms so the information flows properly over the two networks.

Considering all that, the contributions of this work are as follows: (i) a novel CNN architecture, called DS-FNet, that combines a modified U-Net semantic segmentation model with a boundary detection model via attention-aware mechanisms, and (ii) a novel multi-stain data set, named WSI_Fiocruz, with fine annotation of glomeruli from a larger number of renal WSIs of humans.

3. Glomeruli, stains and segmentation

Fig. 2Ⓐ illustrates the main components of a nephron – the functional kidney unit –, which is responsible for the blood filtration and exchange of small molecules generating the urine. Each nephron is composed of a glomerulus and its related tubules (**T**), interstitium (**I**), and blood vessels. Even though the distinct disease can affect one or more of the nephron components, the glomerulus is the main target of the lesion in up to 90% of the renal diseases that require kidney biopsy (Gesualdo et al., 2004; Polito et al., 2010; Dos-Santos et al., 2017).

The afferent arteriole (**Art**) is a final branch of the renal arteries (**Ar**), giving rise to the glomerulus, which is an almost spherical structure, formed by entangled vascular capillary tufts. As transiting through these capillaries, the blood is filtered across their walls. The liquid filtrated from blood is submitted to many exchanges of ions and small molecules, in the renal tubules. This process leads to the production of approximately two liters of urine per day, containing metabolic waste. The constituents of the glomerular filter are the endothelial cell (**EnC**), the underlying glomerular basement membrane, and the visceral epithelial cell (**EpC**) (podocytes (**P**)). The glomerular capillary tufts are supported by a central axis (mesangium) of mesangial cells (**Mes**) immersed in a mesangial matrix. the glomerulus is delimited by the Bowman's capsule (**BC**). It is composed by a fibrous tissue internally coated by a simple squamous epithelium layer. The BC encloses a space called Bowman's space (**BS**), located between the parietal and visceral epithelial cell layers.

Table 1
Summary of the state-of-the-art on kidney tissue segmentation.

Reference	#WSI	#Glom.	Method	Stain
de Bel et al. (2018)	15	944	FCN M-FCN U-Net	PAS
Marsh et al. (2018)	48	3867	CNN(VGG-16) LoG Blob-detection Linear Regression	HE
Gadermayr et al. (2019)	24	–	SW-CNN U-Net-S U-Net-D	PAS
Kannan et al. (2019)	275	745	Inception V3 Heatmap Otsu binarization Distance transform Watershed segmentation	TRI
Lutnick et al. (2019)	21	1147	DeepLabv2	HE PAS
Ginley et al. (2019)	79	616	DeepLabv2 RNN(LSTM) Color deconvolution Color space transformation Otsu binarization Naive Bayesian	PAS
Ginley et al. (2020)	65	393	DeepLabv2	PAS
Bueno et al. (2020)	47	1245	U-Net (SegNet AlexNet)	PAS
Altini et al. (2020b)	26	2772	SegNet DeepLabv3+	PAS
Mei et al. (2020)	–	819	Unet Feature Maps	TRI PAMS
Zeng et al. (2020)	400	12,418	ARPS (U-Net DenseNet LSTM-GCNet 2D V-Net)	PAS
Altini et al. (2020a)	26	2772	Fast-RCNN Mask-RCNN	PAS
Jha et al. (2021)	61	1334	Mask-RCNN U-Net DeepLabv3	HE PAS PAMS
Jayapandian et al. (2021)	459	1196	U-Net	HE PAS PAMS TRI
Bouteldja et al. (2021)	168	2611	Modified U-Net	PAS
Jiang et al. (2021)	348	8665	Cascade Mask R-CNN	PAS PAMS TRI
Davis et al. (2021)	258	24,133	U-Net	HE
Hermsen et al. (2019)	60	238	U-Net	PAS
Ours	**660**	**5309**	**DS-FNet**	**HE** **PAS** **PAMS** **TRI**

Pre-analytical processing of renal biopsies is a fundamental step of a proper morphological analysis. This step includes sample fixation, paraffin-embedding, sectioning in 2–3 micrometers thin slices, and staining by using different techniques. Most of the nephropathology laboratories use a set of four staining techniques in diagnostic routine: Periodic acid-Schiff (PAS), periodic acid-methenamine sil-ver (PAMS), hematoxylin-eosin (HE), and a trichrome (TRI) stain. These staining techniques use different pigments and set the conditions to attach them

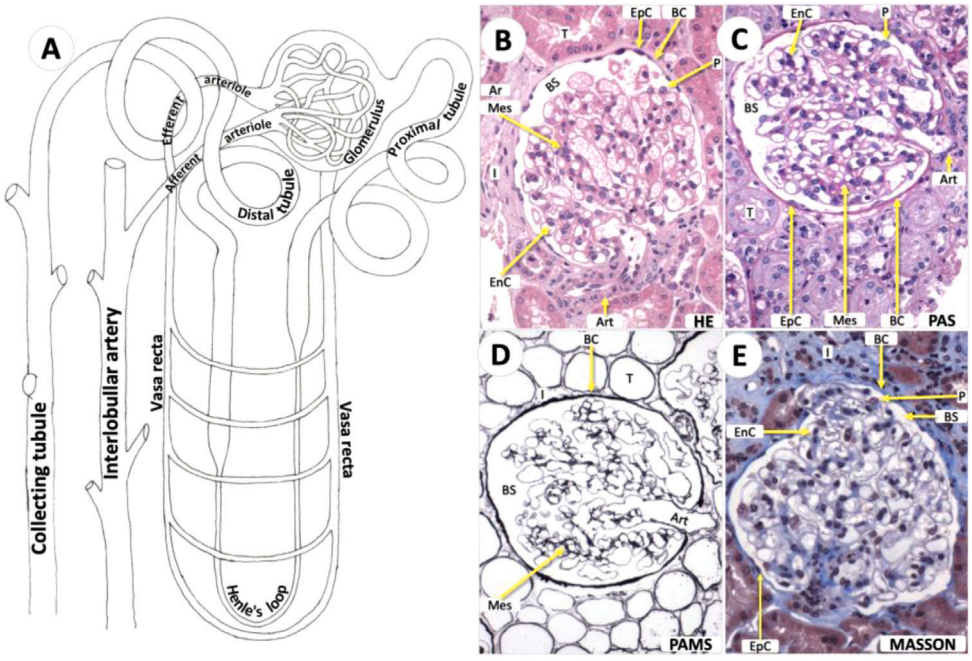


Fig. 2. Schematic and histological view of the nephron and examples of different stained glomerulus images. Ⓐ - Graphic representation of the nephron: Blood flows from the **interlobullar artery** to the **afferent arteriole**, enter the glomerular capillaries and leaves the glomerulus through the **efferent arteriole**. The blood is filtrated in the glomerulus and the filtrate flows through the renal tubules where solutes are exchanged; urine is concentrated and discharged through the renal **collecting tubule**. Ⓑ – **HE** stain reveals most of the microscopic structures such as cells, and supporting matrix. Ⓒ – **PAS** stain highlights the supporting membranes such as Bowman's capsule and mesangium. Ⓓ – **PAMS** is a silver stain that enhance visualization of extracellular matrix (membrane and part of interstitial components). Ⓔ – **Masson's trichrome** stain highlights interstitial extracellular matrix. The main structures highlighted in the stained glomeruli are: **Ar** = artery, **Art** = arteriole, **BC** = Bowman's capsule, **BS** = Bowman's space, **EnC** = endothelial cell, **EpC** = parietal epithelial cell, **I** = interstitium, **Mes** = mesangium, **P** = podocyte, and **T** = tubule.

to different tissue structures. The attachment occurs according through chemical and physical properties of components of these structures. For instance, during HE staining (see Fig. 2Ⓑ) hematoxylin binds to anionic components of the cell nucleus while eosin binds to cathionic components of the cell cytoplasm. Therefore, the use of different staining techniques allows a proper visualization of different tissue structures. In the study of renal biopsies, HE is used to obtain a general view of the tissue. Cell nucleus and cytoplasm are clearly stained contrasting with the light staining of the extracellular matrix. PAS (see Fig. 2Ⓒ) highlights basement membranes and sugar aggregates. There are a variety of silver stain techniques such as Jones methenamine silver and PAMS (see Fig. 2Ⓓ) that clearly delineates details of normal or altered basement membranes. TRI, such as Masson (see Fig. 2Ⓔ) or Azan trichromes, stains collagen and other extracellular matrix components allowing visualization of cell-extracellular matrix relationship (Cathro et al., 2018; Chang et al., 2012).

In the analysis of renal biopsy, each glomerulus must be identified and individually examined. Renal biopsy analysis requires a trained pathologist and is very time-consuming. Accuracy in glomeruli identification is essential for accessing activity or chronicity of renal diseases.

4. One-to-many-stain glomerulus segmentation

In Fig. 1, the same glomerulus is stained with PAS, PAMS, and HE. Although each stain highlights different structural components, it is possible to recognize the same shape of the glomeruli with all stains. The features highlighted by these stains allow ML models to create different embeddings that represent the glomerulus in different ways. We hypothesize that a semantic segmentation model trained in one coloration, *e.g.* PAS, could find similar features when applied to segment glomeruli in other colors, such as HE, as long as the correct information could pass throughout the higher layer in the networks. Bearing this in mind, we conceived the DS-FNet segmentation model grounded in attention-aware mechanisms.

The conception of DS-FNet followed three main steps. The **first step** consisted of identifying a robust baseline model for the glomeruli segmentation task. For model selection, we considered which model reached the most competitive results reported in the literature for glomerulus segmentation. The baseline segmentation model that met this requirement was the U-Net (Ronneberger et al., 2015). Modified versions of U-Net reached the first places on HubMAP challenge, and this network is also used in the majority of the works in Table 2, as discussed in Section 2.

The U-Net network is based on an encoder–decoder architecture that combines low-level features (edges) with more high-level semantic features. In the **second step**, we modified U-Net by adding attention on the feature channels and on spatial location in order to deal with different edge information present in different stained images. In the original version of U-Net, a VGG (Simonyan and Zisserman, 2014) network is used as an encoder. Instead, we used an EfficientNet B1 (Tan and Le, 2019) as the encoder due to its high balance among accuracy, model size, number of parameters and the inference time. From this version of U-Net, we propose the so-called AU-Net, which is based on attention mechanisms.

In the **last step**, we added a stream to AU-Net in charge of detecting the semantic boundary of the glomerulus. The rationale for adding a semantic boundary detector is that the boundary is one of the main information preserved across different stains. This way, we propose a CNN architecture that contains a semantic boundary detection module ultimately combined with the flow that performs the segmentation conceived in the first and second steps. The combination of segmentation and boundary detection tasks in DS-FNet was essential to achieve high performance on data sets gathered with images of different characteristics.

4.1. AU-Net

Attention modules try to capture a broader context in the image when compared to the image representation captured by convolutional

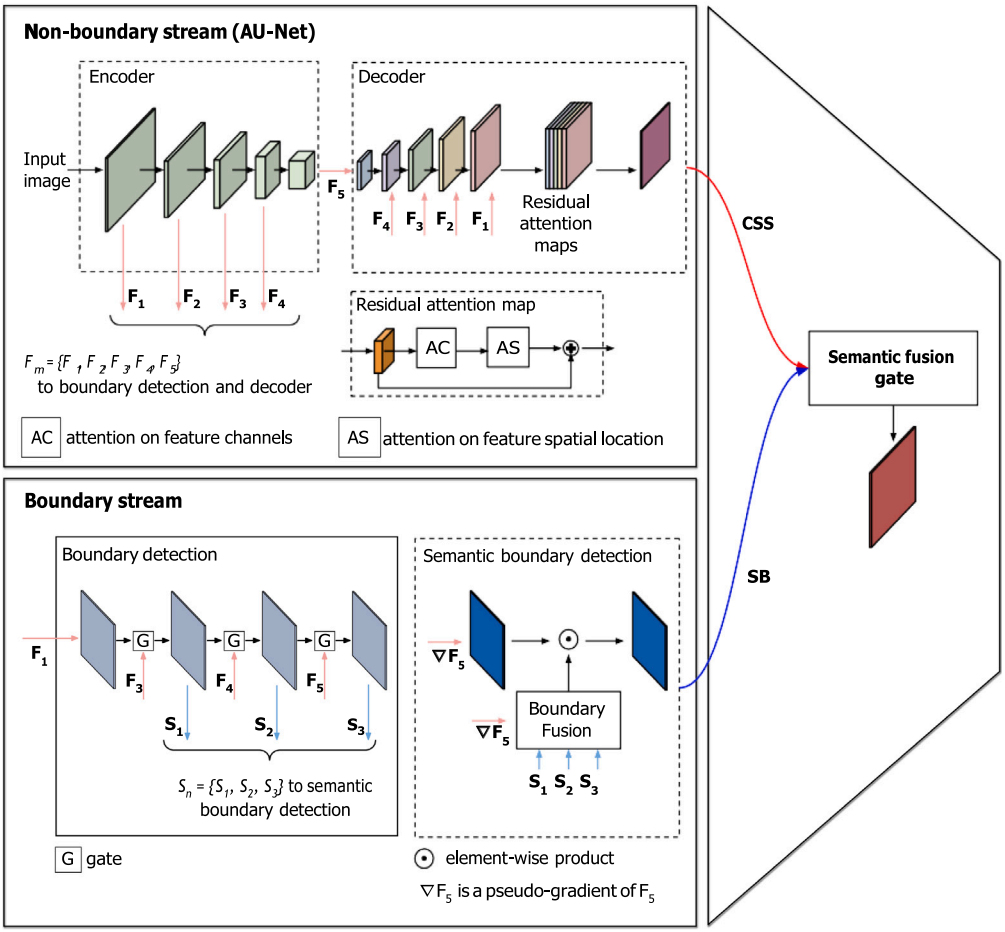


Fig. 3. Detailed architecture of our DS-FNet. The input image is fed to a backbone that generates a set of F_m features maps: F_1, F_2, F_3, F_4, and F_5 with different scales. The feature maps F_n are used in the boundary stream to obtain a set of S_n features: S_1, S_2, and S_3. Each S_n in the boundary stream represents the boundary from the feature maps of different scales from F_n. ∇F_5 is a pseudo gradient obtained using a maxpool with kernel 2 and stride 3 over F_5. The semantic boundary detection module combines ∇F_5 and S_1, S_2, and S_3 to obtain the semantic boundaries (**SE**), which ultimately combine with the coarse semantic segmentation (**CSS**) through the attention gate to output the segmented image. The residual attention maps is the module that sequentially combines the attention on feature channel (**AC**) and on feature spatial location (**AS**).

layers. In image segmentation and classification tasks, the attention module has been used to improve results in both aforementioned tasks (Huang et al., 2019; Zhang et al., 2020; Park et al., 2018; Hu et al., 2018). In Hu et al. (2018), a simple attention module per feature channel is used to improve the performance of a classification network. Here we explored the attention module proposed by Hu et al. (2018) by modifying it to work with feature spatial attention. The rationale to explore the attention module in U-Net was based on increasing the context representation of the local convolutional features, thus improving AU-Net output to get a more robust semantic segmentation. Attention module on feature channels, $\tilde{\mathbf{A}}$, is given by

$$\tilde{\mathbf{A}} = \mathbf{A} \odot \sigma(\mathrm{fc}(\Psi(\mathbf{A}))), \tag{1}$$

where $\mathbf{A}$ represents the input features, fc denotes the fully connected network, $\odot$ denotes an element-wise product, and Ψ is a global average pooling function.

In turn, attention module on feature spatial location, $\tilde{\mathbf{E}}$, is given by

$$\tilde{\mathbf{E}} = \mathbf{E} \odot \sigma(\circledast_{1\times1}(\circledast_{1\times1}(\mathbf{E}))), \tag{2}$$

where $\mathbf{E}$ represents the input features, σ is a sigmoid function, $\odot$ denotes a element-wise product, and $\circledast_{1\times1}$ represent a convolution operation with filters of size 1×1, with zero-padding and stride of 1.

Attention on feature channels and feature spatial location modules are sequentially combined in the **residual attention map** (RAM) module, as shown in Fig. 3, in the non-boundary stream module. The RAM is defined as

$$\mathrm{RAM(D)} = \tilde{\mathbf{E}}(\tilde{\mathbf{A}}(\circledast_{1\times1}(\mathrm{D}))) + \mathrm{D}, \tag{3}$$

where D is an array of features.

After each block of the U-Net decoder, we apply the RAM module, obtaining five feature matrices. The matrices are concatenated and convoluted to obtain the final semantic segmentation in the top of AU-Net.

4.2. DS-FNet: Semantic fusion of AU-Net and boundary detection

Fig. 3 depicts our DS-FNet architecture. We combined AU-Net, a non-boundary stream, with a boundary stream via a semantic fusion gate. The boundary stream is composed by a boundary detection module and a semantic boundary detection module. By combining these streams, both segmentation and boundary detection tasks could be optimized, improving the overall segmentation result.

Boundary detection. We use features from different stages of the encoder module of AU-Net, namely, F_1, F_3, F_4, F_5[1] (refer to Fig. 3). These features are sequentially combined by using attention gates (G_A) followed by a residual module and a 1×1 convolution operation. The

[1] Feature F_2 was not used in the boundary stream because it keeps a large similarity with feature F_1.

purpose of each gate, G_A, is to ensure that the information about the boundaries is selected, preventing the non-boundary and boundary information getting mixed. This way, the attention gates (G_A)[2] are denoted as

$$\mathrm{G_A} = \circledast_{1\times1}(F_n \odot (\sigma(\circledast_{1\times1}(F_n \parallel F'_n)) + 1)), \tag{4}$$

where F_n represents the backbone features from the n_{th} stage of AU-Net encoder, $\mathbf{F}'_n$ are the same $\mathbf{F}_n$, $\parallel$ after convoluted in the boundary detection module, $\circledast_{1\times1}$ is a convolutional layer, σ is the sigmoid function, and $\odot$ is an element-wise product. The output of each gate G_A is a matrix defined as $S_1, S_2, S_3 \in \mathbb{R}^{H\times W\times1}$.

Semantic boundary detection (SBD). This module uses the fusion of high-level (semantic) and low-level features based on the work found in Hu et al. (2019). In addition, we incorporated the ∇F_5 and the matrices S_1, S_2, and S_3 as inputs in this module. After weighing S_1, S_2, and S_3, we perform a 1×1 convolution operation. ∇F_5 is obtained by a pseudo-gradient operation in order to achieve semantic boundaries without causing ambiguity during back-propagation, and is given by

$$\nabla F_5 \approx \sigma(F_5 - \mu_{3\times3}(F_5)), \tag{5}$$

where $\mu_{3\times3}$ represents a maximum-pooling operation in the 2D space using a 3×3 kernel.

The **SBD** takes as inputs ∇F_5 and the matrices S_n. First, ∇F_5 is upsampled by two transposed convolution layers with stride of 8 pixels. The result is then split into k slices. Each slice is concatenated with the S_n sets and linearly combined by a $\circledast_{1\times1}$, called F_{edges}. Finally, the features F_{edges} go through an attention gate SBatt, given by

$$\mathrm{SBatt} = L_{adapt}(\nabla F_5) \odot F_{edges}, \tag{6}$$

where, L_{adapt} is a 1×1 convolution followed by batch normalization and ReLU activation, and $\odot$ denotes an element-wise product. The rationale here is to reinforce the boundaries formed by both low-level and high-level features with more semantic information. Finally, the resulting SBatt is convoluted by a $\circledast_{1\times1}$, resulting in the semantic boundary feature $\mathbf{SB} \in \mathbb{R}^{H\times W\times1}$.

Semantic fusion gate. Given the output from AU-Net ($\mathrm{CSS} \in \mathbb{R}^{H\times W\times1}$) and the output from boundary stream ($\mathrm{SB} \in \mathbb{R}^{H\times W\times1}$), we first feed CSS with 2 convolutional sets. The output of the first set is denoted as CSS′ features, while the second set outputs a mask M. A softmax function is applied in the CSS′, obtaining the segmentation S. The matrix SB is multiplied by $1 - \mathrm{M}$. The result is added with S, getting the final segmentation. The semantic fusion gate avoids ambiguity in classifying boundary pixels, as it combines the appropriate information from each stream.

5. Materials and methods

5.1. Data sets

Our experimental analyses relied upon three data sets whose characteristics are summarized in Table 2 and described next.

5.1.1. HuBMAP - "Hacking the Kidney" data set

Sponsored by the National Institutes of Health (NIH), the human biomolecular atlas program (HuBMAP) managed a glomerulus segmentation competition in the Kaggle (Borner et al., 2021) platform. This challenge consisted of developing a supervised model for glomerulus segmentation given a set of 20 human kidney WSIs. The data set includes 11 fresh frozen and 9 formalin-fixed paraffin-embedded (FFPE) PAS-stained kidney images. Among the WSIs, 15 are used for training and 5 are left for testing. Test images and annotations are not available, and model results can only be verified through Kaggle platform submissions. Fig. 4 illustrates some data set samples.

Table 2
Summary of the data set characteristics.

Dataset	Stain	#WSI	#Glom.	Set	Width/Height (avg. px.)
HuBMap[a]	PAS	15	3568	Train	10k/12k
	PAS	5	–	Test	–
NEPTUNE	PAS	203	428	Test	3k/3k
	PAMS	123	295	Test	3k/3k
	HE	157	339	Test	3k/3k
	TRI	137	325	Test	3k/3k
	Total	**620**	**1387**		
WSI_Fiocruz	PAS	10	136	Test	23k/11k
	PAMS	5	70	Test	21k/9k
	HE	10	148	Test	23k/10k
	Total	**30**	**354**		

[a]Note that the columns #Glom. and Width/Height were not reported as they are not available on the Kaggle platform for the private test set.

5.1.2. NEPTUNE subset

This data set includes renal sections collected from the multicenter nephrotic syndrome study network (NEPTUNE) (Jayapandian et al., 2021), here referred as **NEPTUNE subset** (or **NEPTUNE** for short). This data set provides 620 publicly available WSI crops, which are treated here as WSIs for easy reference. The data set annotations were originally provided by five nephropathologists. As we included annotations for partially-occluded glomeruli, our quantities of WSIs and glomeruli differ from Jayapandian et al. (2021). Our study focused on glomeruli whose annotations included Bowman's space (also following annotations in HuBMAP and WSI_Fiocruz data sets), and then we considered only the glomerular units in this data set rather than glomerular tufts as in Jayapandian et al. (2021). The final NEPTUNE subset we used contains 1387 glomeruli, with each stained image extracted with 5× digital magnification. Fig. 5 illustrates some data set samples.

5.1.3. WSI_Fiocruz data set

Our data set contains images from human kidney biopsies collected in referral nephrology services of public hospitals in Bahia, Brazil. The renal tissues were fixed in formalin-acetic acid-alcohol (FAA). Sections of 2 μm were stained with PAS, PAMS, and HE. The WSI images were captured using a VSI500 Olympus scanner with 40× magnification. The glomeruli were manually annotated by a senior nephropathologist. Fig. 6 illustrates some data set samples.

5.2. Implementation details

Throughout this work, the experiments followed the same training and testing protocol for all architectures. We used the Efficientnet B1 (Tan and Le, 2019) as an encoder for all network models, removing the last dense layer used for classification. We adopted a polynomial learning rate schedule defined as $lr = base * (1 - \frac{iter}{total\_iter})^{power}$, where $base$ is the initial learning rate with a value of 10^{-4}, $iter$ is the number of iterations during training, $total\_iter$ is the total number of iterations, and $power$ is the degree of the polynomial. As an optimization method, we used the parameterized SGD with a momentum of 0.9 and weight decay set to 0.0005.

The experiments were run in a computer equipped with eight NVIDIA V100, each one containing 16 GB of memory. The batch size for training was 16 images. We divided the WSI images into crops of 1024×1024 pixels with a stride of 512 pixels. Each crop was resized to 320×320 and then used for training. We used mixed-precision during training to reduce memory consumption in order to increase training speed.

[2] Unlike an attention module, the attention gate tries to split the feature set into two or more separate feature sets and not only highlights important information for the next block in the network.

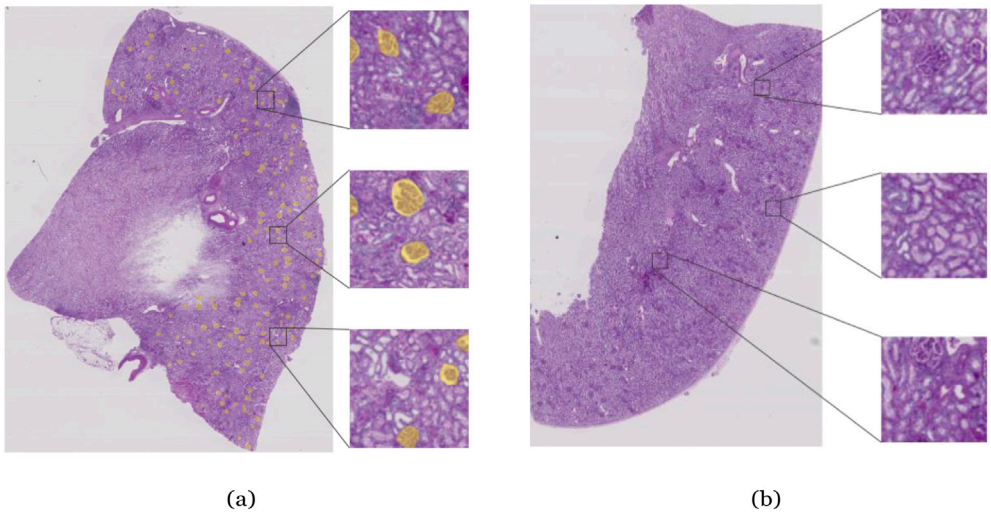


Fig. 4. Image samples from the HuBMAP data set. In (a), an image sample of the train set with the glomeruli annotated in yellow; in (b), an image sample of the testing set.

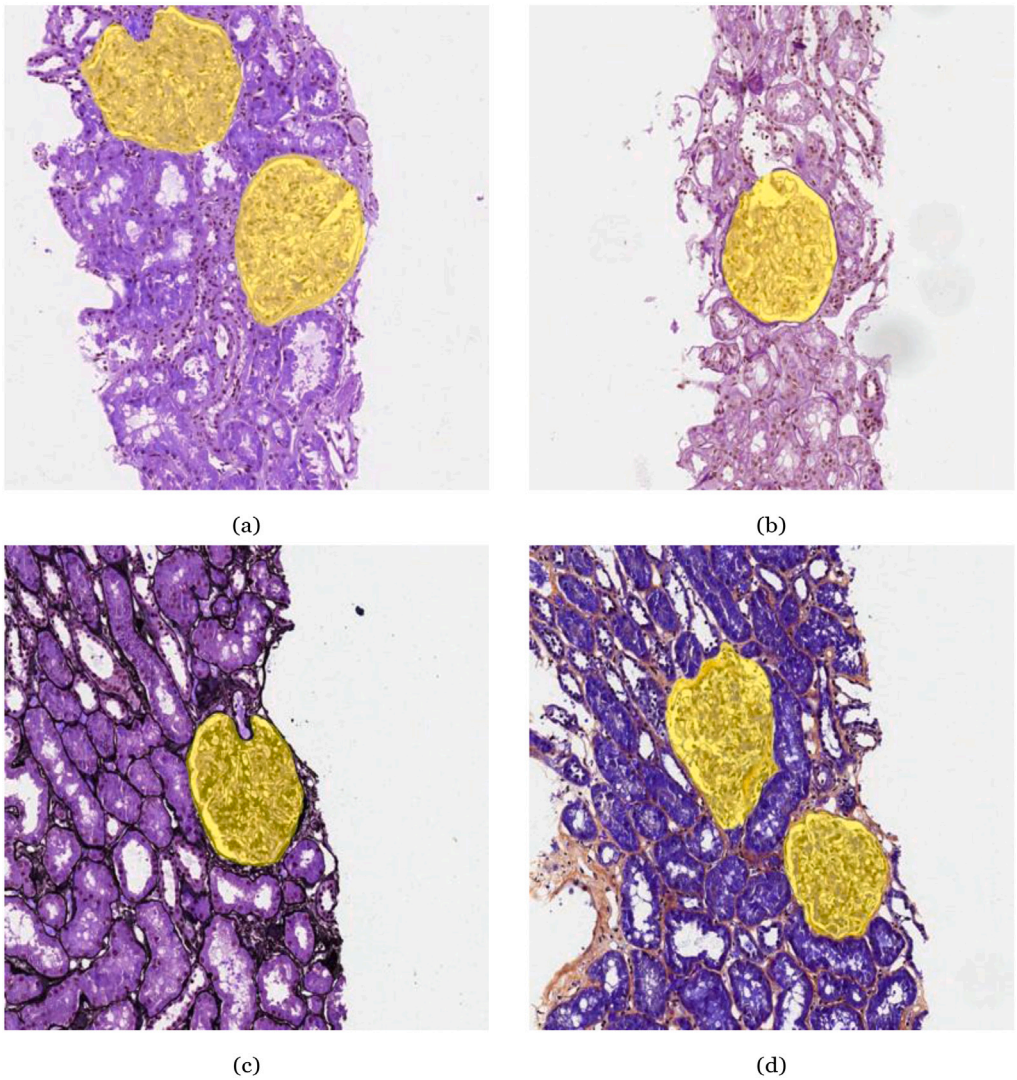


Fig. 5. Image samples from the NEPTUNE data set. (a) HE image sample, (b) PAS image sample, (c) PAMS image sample and (d) TRI image sample.

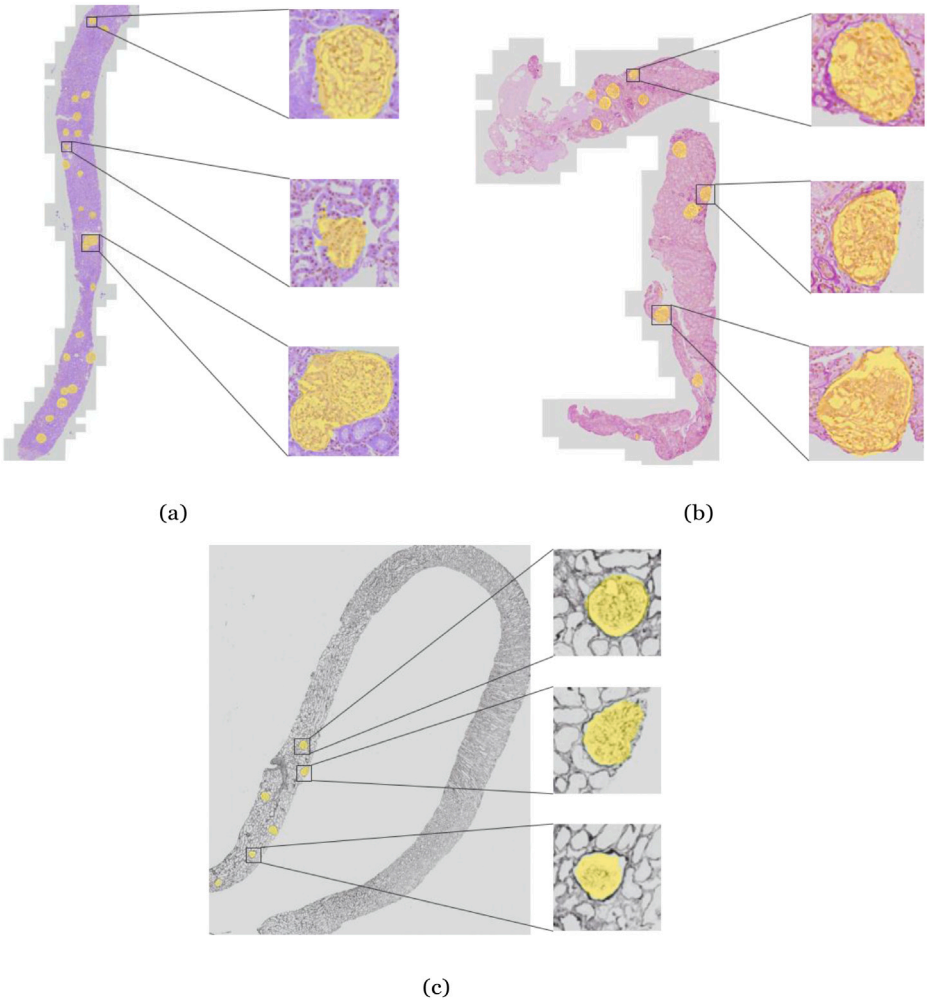


Fig. 6. Image samples from the WSI_Fiocruz data set. (a) HE image sample, (b) PAS image sample, and (c) PAMS image sample.

Training process. To train the models, we split the HubMAP data set into four folds (see Table 2), having three folds for training and one left for validation. After four different training sets, we had four different models according to a given performance metric in the validation set. We used the dice score (DSC) to evaluate whether the validation performance of each epoch has increased or not, keeping only the model whose epoch returned the highest DSC. We set a total of thirty epochs for training each architecture, because after some preliminary experiments the networks best reached convergence. The glomerulus segmentation data sets are unbalanced if we consider the proportion of glomerulus and non-glomerulus pixels. To avoid a considerable unbalanced data presentation to the network, we have divided the patch set into two types: The first set contains patches without pixels of glomerulus, while the second set does contain. Thus, an equivalent amount of each set was selected to be used during the training.

Self-training. We refined the training of all networks by using a self-training approach (Lee et al., 2013) to improve the generalization of the final trained models. For that, it was used the public training set of HubMap. First we trained a model from all networks using just the label images with the initial aforementioned parameters. We then run these models on the images belonging to the public testing set of HubMAP (see Table 2). With the segmentation masks in hand obtained over initial unlabeled images, the predicted images were added to the existing training set, thereby increasing our initial training data. Finally, we retrained all the models considering a learning rate of 10^{-6} in ten epochs with the updated training set (manual- and machine-labeled images).

5.3. Evaluation protocol

To assess the performance of our proposed method, we followed a set of experiments with the goal of firstly define the best parameters of our network, then evaluating our best architecture over the chosen data sets. We conducted an ablation study on DS-FNet considering only the HubMAP data set by means of the Kaggle platform. In the ablation study, we ranked the best model of the original U-Net and AU-Net. These early experiments was mainly performed to assess what would be the best combination of loss functions and if the use of a class supervision[3] would improve the final segmentation performance. In this stage, considering the differences of U-Net and AU-Net, we also evaluated the impact of the feature channel and feature spatial location attention modules. The second set of experiments was to compare the best model obtained in the ablation stage with the other competitors in Kaggle's leaderboard. The last set of experiments was thought to evaluate the generalization of DS-FNet in terms of different stains and data sets: After trained only on the HubMAP data set, our best model was evaluated on the NEPTUNE subset and the WSI_Fiocruz data sets. Given that specific training, validation, and testing splits were not available in the NEPTUNE subset, it was not possible to directly compare our proposed network with the results found in Jayapandian et al. (2021).

[3] Class supervision is a branch for classifying the presence of glomeruli or part of the glomeruli in image patches during training.

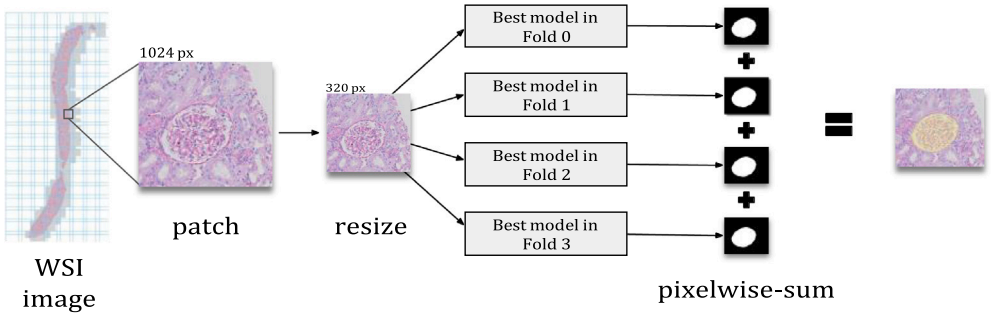


Fig. 7. Ensemble method used in the evaluation protocol. First, a WSI image is divided into patches 1024 × 1024 pixels and each patch is resized to 320 × 320 pixels. Each patch follows for the best model of each fold. The results of the predictions of each best model in each fold are pixel-wise summed to obtain the segmentation mask of each patch. We used the ensemble to obtain the final segmentation for each one of evaluated networks.

Table 3
Summary of the ablation on the HuBMAP private testing data considering U-NET and AU-Net and all possible combinations of loss functions (BCE, dice, WBCE) and class supervision technique.

Method	Loss function		Class	DSC
	BCE	dice	supervision	
U-Net	✓	–	–	0.87
U-Net	✓	✓	–	0.89
U-Net	✓	✓	✓	0.91
AU-Net	✓	–	–	0.90
AU-Net	✓	✓	–	0.91
AU-Net	✓	✓	✓	0.92

Table 4
Results of all networks on the HuBMAP private testing data, using bce and dice loss functions and class supervision.

Method	DSC
U-Net	0.91
AU-Net	0.92
U-Net++ (Zhou et al., 2018)	0.90
U-Net3Plus (Huang et al., 2020)	0.92
ResU-Net (Diakogiannis et al., 2020)	0.89
DeepLabV3+ (Chen et al., 2018)	0.88
DS-FNet	**0.95**

Instead of using just one model of each network architecture, we ensemble the four models obtained through the 4-fold cross-validation procedure described in Section 5.2. Fig. 7 illustrates how the models are ensembled to provide the final segmentation masks. The ensemble segmentation strategy starts with the division of the image into 1024 × 1024 patches considering 512 pixels of stride between patches. Each patch is resized to 320 × 320 pixels, and only then it is used as input to the segmentation models. The results of each model for each corresponding patch are summed and then a sigmoid function is applied to the resulting image. Finally, we resized the mask with the segmentation of the glomerulus to the original input size.

6. Results and discussion

Each architecture listed in this section is a result of the ensemble depicted in Fig. 7. The results of the ablation study on the HuBMAP private testing set are summarized in Table 3. We did not use wbce in U-Net and AU-Net because this loss was designed for boundary detection (Yu et al., 2017). The wbce loss considers equally important classes in each image, even if there are very rare classes in the image. For the pure semantic segmentation task, the use of this loss can be a problem as it can reduce the performance of the segmenter in terms of coverage. The results demonstrated that the introduction of the dice loss optimization allowed slight improvements for both U-Net and AU-Net while the class supervision also improved the AU-Net performance. Considering the best parameters of AU-Net, and a DS-FNet that also uses wbce loss because the aforementioned reasons, our proposed network achieved the best results on the HuBMap testing private set as presented in Table 4.

Table 5
Results of the Kaggle competition on the HuBMAP private testing data set (Borner et al., 2021), considering the four first competitors and DS-FNet. It is noteworthy that we did not participate in the competition, but posteriorly evaluated our method.

Method	DSC
U-Net (SE_ResNext101)	0.9503
U-Net (EfficientNet-B1)	0.9503
DS-FNet	0.9505
No method description	0.9507
U-Net (SE_ResNext101)	**0.9515**

Table 6
Results on the data sets from different medical labs stained only in PAS.

Method	HuBMap	NEPTUNE	WSI_Fiocruz	Avg. DSC
U-Net	0.89	0.89	0.67	0.82
AU-Net	0.92	0.91	0.83	0.89
U-Net++ (Zhou et al., 2018)	0.90	0.89	0.78	0.86
U-Net3Plus (Huang et al., 2020)	0.92	0.90	0.81	0.88
ResU-Net (Diakogiannis et al., 2020)	0.89	0.86	0.71	0.82
DeepLabV3+ (Chen et al., 2018)	0.88	0.89	0.70	0.82
DS-FNet	**0.95**	**0.92**	**0.86**	**0.91**

Also, when the best parameterized DS-FNet is benchmarked with the other methods in the HuBMAP leaderboard in Kaggle platform,[4] it is clear in Table 5 that our network presents similar results to the four best methods (among 1216 submissions). Indeed, considering the four decimal places of the competition, our model is just 0.01 percentage point behind the very first place.

Considering the best DS-FNet on HuBMap data set, our work was to evaluate how robust would be our network: Over data sets from different labs, using only the PAS stain and also images stained with different techniques. Table 6 shows the result of the average DSC of each one of the networks on the three data sets, with images stained only in PAS. DS-FNet reached the best results on the three data sets. Besides outperforming U-Net by 9 percentage points, the DS-FNet also outperformed AU-Net (the second best result) by two percentage points. As these results are taken from images stained in PAS, it was already expected that the differences between the results of the attention-based networks (AU-Net and DS-FNet) were not considerably high in comparison with the other ones.

In Table 7, the results evidenced that DS-FNet consistently outperformed the other methods for all types of stains. Now, higher differences between the network results were found when the stains in

[4] https://www.kaggle.com/c/hubmap-kidney-segmentation/leaderboard/.

Table 7
Results on the data sets from different medical labs stained with periodic acid-methenamine silver (PAMS), hematoxylin-eosin (HE), and trichrome (TRI).

Dataset	Method	HE	PAMS	TRI	Avg. DSC
NEPTUNE	U-Net	0.68	0.81	0.66	0.72
	AU-Net	0.72	0.86	0.75	0.78
	U-Net++ (Zhou et al., 2018)	0.67	0.83	0.65	0.72
	U-Net3Plus (Huang et al., 2020)	0.71	0.85	0.76	0.77
	ResU-Net (Diakogiannis et al., 2020)	0.65	0.80	0.63	0.69
	DeepLabV3+ (Chen et al., 2018)	0.66	0.81	0.67	0.71
	DS-FNet	**0.83**	**0.88**	**0.80**	**0.84**
WSI_Fiocruz	U-Net	0.67	0.63	–	0.65
	AU-Net	0.67	0.71	–	0.69
	U-Net++ (Zhou et al., 2018)	0.67	0.65	–	0.66
	U-Net3Plus (Huang et al., 2020)	0.69	0.71	–	0.70
	ResU-Net (Diakogiannis et al., 2020)	0.63	0.61	–	0.62
	DeepLabV3+ (Chen et al., 2018)	0.70	0.67	–	0.69
	DS-FNet	**0.80**	**0.79**	–	**0.80**

the prediction were different from the stain in the training. Even though the average results in that table is lower than those in Table 6, it is clear that DS-FNet still preserves considerably high DSC, while substantially increases the average difference from the U-Net3Plus (second best result) to 7 and 10 percentage points in the Neptune subset and the WSI_Fiocruz data set, respectively. These findings corroborate with the initial hypothesis of generalization potential of the proposed method in the challenging simultaneous cross-lab and cross-stain segmentation tasks.

Over the results in Tables 6 and 7, we applied the McNemar's test to evaluate whether there was a statistically significant difference between the results of the DS-FNet and the other networks presented in those tables. As the granularity of the segmentation is in the pixel level, the semantic segmentation of the pixel was used as the classification result for the statistical test. By taking the p-value as lower than 0.05, the test showed that DS-FNet results were statistically different from the networks tested in Tables 6 and 7, for the NEPTUNE and WSI_Fiocruz data sets.

6.1. Qualitative results

Figs. 8 and 9 illustrate visual comparisons of the segmentation results of all networks over challenging image samples of the testing data sets. In Fig. 8, on the NEPTUNE subset samples, the AU-Net and DS-FNet were able to reduce the false positives (yellow) significantly. DS-FNet was also effective in significantly reducing the false negatives (red) by including some glomeruli regions completely missed by the other methods. On the WSI_Fiocruz data set, AU-Net better identified positive regions at the expense of too many false positives, and the DS-FNet again outperformed it. In the samples shown in Fig. 9, visual results of the DS-FNet are compared with U-Net++ (Zhou et al., 2018), U-Net3Plus (Huang et al., 2020), ResU-Net (Diakogiannis et al., 2020), and DeepLabV3+ (Chen et al., 2017). Considering those challenging samples, DS-FNet is significantly superior to the other networks concerning true positive results (in green), presenting few false negatives (in red). The only situations in which the other networks showed results close to the DS-FNet in the visual inspection of Fig. 9 were those found in PAS and PAMS on NEPTUNE subset, in detriment of presenting some false negatives. The DS-FNet presented near-perfect segmentations in these particular samples, especially on the boundaries.

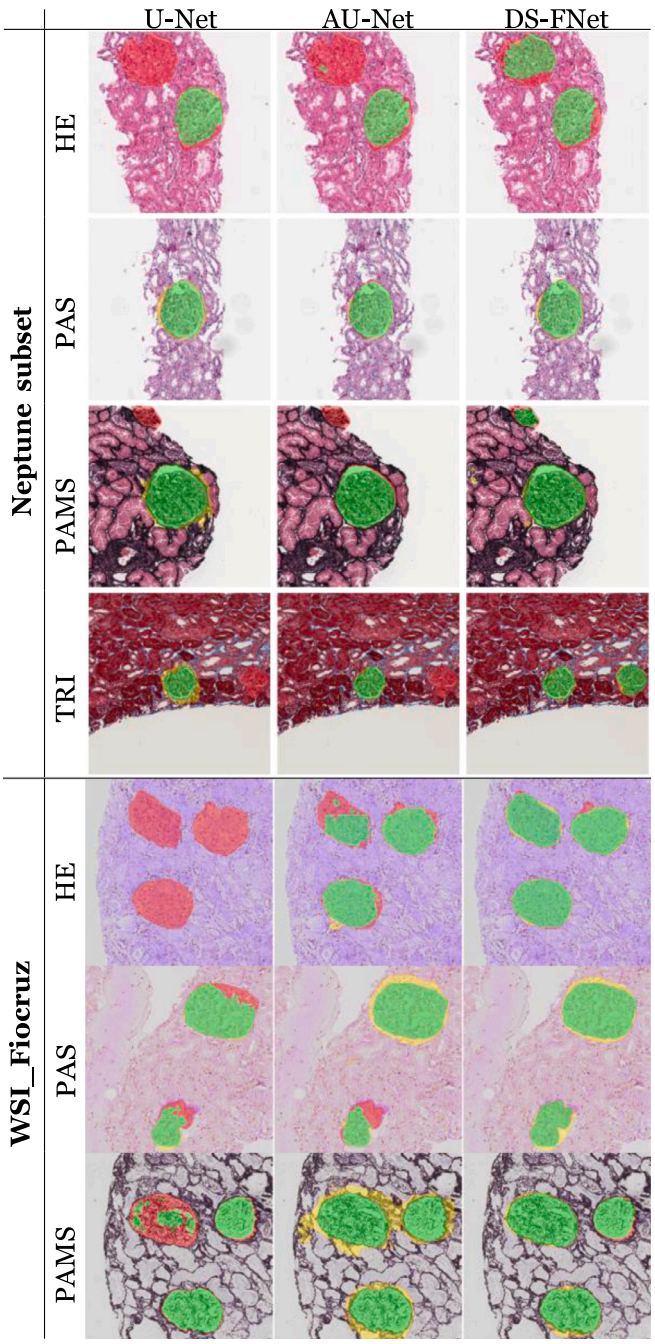


Fig. 8. Some results of U-Net, AU-Net, DS-FNet on NEPTUNE subset (Jayapandian et al., 2021) and our WSI_Fiocruz data sets using four staining techniques: Periodic acid-Schiff (PAS), periodic acid-methenamine silver (PAMS), hematoxylin-eosin (HE), and trichrome (TRI). Glomeruli in yellow are false positives, in red are false negatives, and in green are true positive.

7. Conclusion

Our work brings as contributions a novel DL approach evaluated via a challenging protocol. Our proposed network called DS-FNet explored the possibility of using boundary detection combined with a semantic segmentation network to improve one-to-many-stain segmentation. The empirical evidence presented in this work shows that our proposed network has a superior result when compared with six other DL networks in all evaluations over data sets characterized by cross-stain and cross-lab scenarios. The experimental results also followed the theoretical expectation that glomeruli stained with different techniques keep common characteristics, and this was definitely explored by the attention mechanisms precisely placed throughout our network to integrate non-boundary and boundary information. As future work, we are evaluating DS-FNet over other histological renal structures. We also plan to expand the number of WSIs to 400 in the WSI_Fiocruz data set,

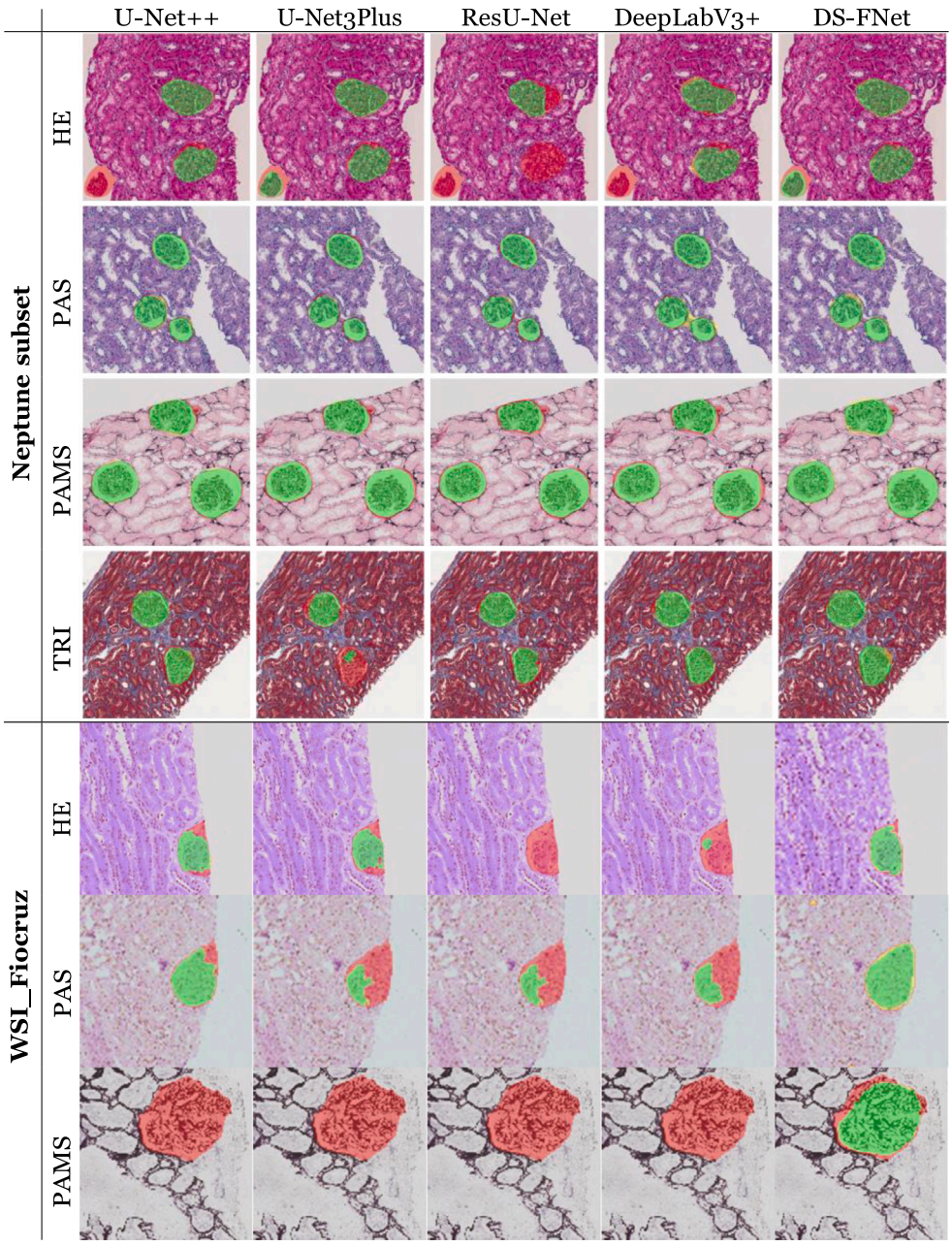


Fig. 9. Some results of U-Net++ (Zhou et al., 2018), U-Net3Plus (Huang et al., 2020), ResU-Net (Diakogiannis et al., 2020), DeepLabV3+ (Chen et al., 2017), and DS-FNet on NEPTUNE subset (Jayapandian et al., 2021) and our WSI_Fiocruz data set using four staining techniques: Periodic acid-Schiff (PAS), periodic acid-methenamine silver (PAMS), hematoxylin-eosin (HE), and trichrome (TRI). Glomeruli in yellow are false positives, in red are false negatives, and in green are true positive.

making it available with the goal of providing a challenging data set to the researchers in this field.

Ethical considerations

This work was conducted in accordance with resolution No. 466/12 of the Brazilian National Health Council. To preserve confidentiality, the images (including those shown in the paper) were separated from other patient's data. No data presented herein allows patient identification. All the procedures were approved by the Ethics Committee for Research Involving Human Subjects of the Gonçalo Moniz Institute from the Oswaldo Cruz Foundation (CPqGM/FIOCRUZ), protocols No. 188/09 and No. 1.817.574.

CRediT authorship contribution statement

Jefferson Silva: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Visualization, Data curation. **Luiz Souza:** Writing – original draft, Visualization, Data curation. **Paulo Chagas:** Formal analysis, Writing – original draft, Writing – review & editing. **Rodrigo Calumby:** Formal analysis, Writing – original draft, Writing – review & editing. **Bianca Souza:** Data curation. **Izabelle Pontes:** Data curation. **Angelo Duarte:** Investigation, Writing – review & editing, Supervision, Project administration. **Nathanael Pinheiro:** Writing – original draft. **Washington Santos:** Investigation, Preparation, Writing – original draft, Writing – review & editing, Supervision, Project administration. **Luciano Oliveira:** Investigation, Methodology, Preparation, Writing – original draft, Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The PathoSpotter project is partially sponsored by the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB), Brazil, grants TO-P0008/15 and TO-SUS0031/2018, and by the Inova FIOCRUZ grant. Washington Santos and Luciano Oliveira have research scholarships from Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil, grants 306779/2017 and 308580/2021-4, respectively. Luiz Otávio and Paulo Chagas have scholarships from FAPESB, Brazil, grants TO-BOL0660/2018 and TO-BOL0344/2018, respectively

References

Altini, N., Cascarano, G.D., Brunetti, A., De Feudis, I., Buongiorno, D., Rossini, M., Pesce, F., Gesualdo, L., Bevilacqua, V., 2020a. A deep learning instance segmentation approach for global glomerulosclerosis assessment in donor kidney biopsies. Electronics 9 (11), 1–21.

Altini, N., Cascarano, G.D., Brunetti, A., Marino, F., Rocchetti, M.T., Matino, S., Venere, U., Rossini, M., Pesce, F., Gesualdo, L., et al., 2020b. Semantic segmentation framework for glomeruli detection and classification in kidney histological sections. Electronics 9 (3), 1–15.

Alturkistani, H.A., Tashkandi, F.M., Mohammedsaleh, Z.M., 2016. Histological stains: a literature review and case study. Glob. J. Health Sci. 8 (3), 72–79.

Bancroft, J.D., Gamble, M., 2008. Theory and Practice of Histological Techniques. Elsevier health sciences.

Bandari, J., Fuller, T.W., Turner, R.M., D'Agostino, L.A., 2016. Renal biopsy for medical renal disease: indications and contraindications. Can. J. Urol. 23 (1), 8121–8126.

Barisoni, L., Gimpel, C., Kain, R., Laurinavicius, A., Bueno, G., Zeng, C., Liu, Z., Schaefer, F., Kretzler, M., Holzman, L.B., et al., 2017. Digital pathology imaging as a novel platform for standardization and globalization of quantitative nephropathology. Clin. Kidney J. 10 (2), 176–187.

Bellur, S.S., Roberts, I.S., Troyanov, S., Royal, V., Coppo, R., Cook, H.T., Cattran, D., Arce Terroba, Y., Asunis, A.M., Bajema, I., et al., 2019. Reproducibility of the oxford classification of immunoglobulin A nephropathy, impact of biopsy scoring on treatment allocation and clinical relevance of disagreements: evidence from the VALidation of IGA study cohort. Nephrol. Dial. Transplant. 34 (10), 1681–1690.

Borner, K., deSouza, A., Vuppalanchi, D., Holland, R., Allam, S., Drury, C., 2021. Hubmap - hacking the kidney. URL https://www.kaggle.com/c/hubmap-kidney-segmentation.

Bouteldja, N., Klinkhammer, B.M., Bülow, R.D., Droste, P., Otten, S.W., von Stillfried, S.F., Moellmann, J., Sheehan, S.M., Korstanje, R., Menzel, S., et al., 2021. Deep learning–based segmentation and quantification in experimental kidney histopathology. J. Am. Soc. Nephrol. 32 (1), 52–68.

Bueno, G., Fernandez-Carrobles, M.M., Gonzalez-Lopez, L., Deniz, O., 2020. Glomerulosclerosis identification in whole slide images using semantic segmentation. Comput. Methods Programs Biomed. 184, 1–10.

Cathro, H.P., Shen, S.S., Truong, L.D., 2018. Diagnostic histochemistry in medical diseases of the kidney. In: Seminars in Diagnostic Pathology, Vol. 35. Elsevier, pp. 360–369.

Chang, A., Gibson, I.W., Cohen, A.H., Weening, J.W., Jennette, J.C., Fogo, A.B., 2012. A position paper on standardizing the nonneoplastic kidney biopsy report. Hum. Pathol. 43 (8), 1192–1196.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Trans. Pattern Anal. Mach. Intell. 40 (4), 834–848.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: ECCV.

Davis, R.C., Li, X., Xu, Y., Wang, Z., Souma, N., Sotolongo, G., Bell, J., Ellis, M., Howell, D., Shen, X., et al., 2021. Deep learning segmentation of glomeruli on kidney donor frozen sections. pp. 1–27, medRxiv.

de Bel, T., Hermsen, M., Smeets, B., Hilbrands, L., van der Laak, J., Litjens, G., 2018. Automatic segmentation of histopathological slides of renal tissue using deep learning. In: Tomaszewski, J.E., Gurcan, M.N. (Eds.), Medical Imaging 2018: Digital Pathology, Vol. 10581. International Society for Optics and Photonics, SPIE, pp. 285–290. http://dx.doi.org/10.1117/12.2293717.

Diakogiannis, F.I., Waldner, F., Caccetta, P., Wu, C., 2020. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS J. Photogramm. Remote Sens. 162, 94–114.

Dos-Santos, W.L.C., Sweet, G.M.M.a., Azevêdo, L.G., Tavares, M.B.a., Soares, M.F.S., Melo, C.V.B.d., Carneiro, M.F.M., Santos, R.F.d.S., Conrado, M.C., Braga, D.T.L., et al., 2017. Current distribution pattern of biopsy-proven glomerular disease in Salvador, Brazil, 40 years after an initial assessment. Braz. J. Nephrol. 39, 376–383.

Gadermayr, M., Dombrowski, A.-K., Klinkhammer, B.M., Boor, P., Merhof, D., 2019. CNN cascades for segmenting sparse objects in gigapixel whole slide images. Comput. Med. Imaging Graph. 71, 40–48.

Gallego, J., Swiderska-Chadaj, Z., Markiewicz, T., Yamashita, M., Gabaldon, M.A., Gertych, A., 2021. A U-Net based framework to quantify glomerulosclerosis in digitized PAS and H&E stained human tissues. Comput. Med. Imaging Graph. 89, 101865.

Gesualdo, L., Di Palma, A.M., Morrone, L.F., Strippoli, G.F., Schena, F.P., et al., 2004. The Italian experience of the national registry of renal biopsies. Kidney Int. 66 (3), 890–894.

Ginley, B., Jen, K.-Y., Rosenberg, A., Yen, F., Jain, S., Fogo, A., Sarder, P., 2020. Neural network segmentation of interstitial fibrosis, tubular atrophy, and glomerulosclerosis in renal biopsies. pp. 1–31, arXiv preprint arXiv:2002.12868.

Ginley, B., Lutnick, B., Jen, K.-Y., Fogo, A.B., Jain, S., Rosenberg, A., Walavalkar, V., Wilding, G., Tomaszewski, J.E., Yacoub, R., et al., 2019. Computational segmentation and classification of diabetic glomerulosclerosis. J. Am. Soc. Nephrol. 30 (10), 1953–1967.

Guo, Z., Liu, H., Ni, H., Wang, X., Su, M., Guo, W., Wang, K., Jiang, T., Qian, Y., 2019. A fast and refined cancer regions segmentation framework in whole-slide breast pathological images. Sci. Rep. 9 (1), 1–10.

Hermsen, M., de Bel, T., Den Boer, M., Steenbergen, E.J., Kers, J., Florquin, S., Roelofs, J.J., Stegall, M.D., Alexander, M.P., Smith, B.H., et al., 2019. Deep learning–based histopathologic assessment of kidney tissue. J. Am. Soc. Nephrol. 30 (10), 1968–1979.

Hu, Y., Chen, Y., Li, X., Feng, J., 2019. Dynamic feature fusion for semantic edge detection. In: Proceedings of the 28th International Joint Conference on Artificial Intelligence. In: IJCAI'19, pp. 782–788.

Hu, J., Shen, L., Sun, G., 2018. Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7132–7141.

Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.-W., Wu, J., 2020. Unet 3+: A full-scale connected unet for medical image segmentation. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 1055–1059.

Huang, Z., Wang, X., Huang, L., Huang, C., Wei, Y., Liu, W., 2019. Ccnet: Criss-cross attention for semantic segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 603–612.

Jayapandian, C.P., Chen, Y., Janowczyk, A.R., Palmer, M.B., Cassol, C.A., Sekulic, M., Hodgin, J.B., Zee, J., Hewitt, S.M., O'Toole, J., Toro, P., Sedor, J.R., Barisoni, L., Madabhushi, A., Sedor, J., Dell, K., Schachere, M., Negrey, J., Lemley, K., Lim, E., Srivastava, T., Garrett, A., Sethna, C., Laurent, K., Appel, G., Toledo, M., Barisoni, L., Greenbaum, L., Wang, C., Kang, C., Adler, S., Nast, C., LaPage, J., Stroger, J.H., Athavale, A., Itteera, M., Neu, A., Boynton, S., Fervenza, F., Hogan, M., Lieske, J., Chernitskiy, V., Kaskel, F., Kumar, N., Flynn, P., Kopp, J., Blake, J., Trachtman, H., Zhdanova, O., Modersitzki, F., Vento, S., Lafayette, R., Mehta, K., Gadegbeku, C., Johnstone, D., Quinn-Boyle, S., Cattran, D., Hladunewich, M., Reich, H., Ling, P., Romano, M., Fornoni, A., Bidot, C., Kretzler, M., Gipson, D., Williams, A., LaVigne, J., Derebail, V., Gibson, K., Froment, A., Grubbs, S., Holzman, L., Meyers, K., Kallem, K., Lalli, J., Sambandam, K., Wang, Z., Rogers, M., Jefferson, A., Hingorani, S., Tuttle, K., Bray, M., Kelton, M., Cooper, A., Freedman, B., Lin, J., 2021. Development and evaluation of deep learning–based segmentation of histologic structures in the kidney cortex with multiple histologic stains. Kidney Int. 99 (1), 86–101. http://dx.doi.org/10.1016/j.kint.2020.07.044, URL https://www.sciencedirect.com/science/article/pii/S0085253820309625.

Jha, A., Yang, H., Deng, R., Kapp, M.E., Fogo, A.B., Huo, Y., 2021. Instance segmentation for whole slide imaging: end-to-end or detect-then-segment. J. Med. Imaging 8 (1), 1–16. http://dx.doi.org/10.1117/1.JMI.8.1.014001.

Jiang, L., Chen, W., Dong, B., Mei, K., Zhu, C., Liu, J., Cai, M., Yan, Y., Wang, G., Zuo, L., et al., 2021. A deep learning-based approach for glomeruli instance segmentation from multistained renal biopsy pathologic images. Am. J. Pathol. 191 (8), 1431–1441.

Kannan, S., Morgan, L.A., Liang, B., Cheung, M.G., Lin, C.Q., Mun, D., Nader, R.G., Belghasem, M.E., Henderson, J.M., Francis, J.M., et al., 2019. Segmentation of glomeruli within trichrome images using deep learning. Kidney Int. Rep. 4 (7), 955–962.

Lee, D.-H., et al., 2013. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In: Workshop on Challenges in Representation Learning, ICML, Vol. 3. (2), p. 896.

Lutnick, B., Ginley, B., Govind, D., McGarry, S.D., LaViolette, P.S., Yacoub, R., Jain, S., Tomaszewski, J.E., Jen, K.-Y., Sarder, P., 2019. An integrated iterative annotation technique for easing neural network training in medical image analysis. Nat. Mach. Intell. 1 (2), 112–119.

Marsh, J.N., Matlock, M.K., Kudose, S., Liu, T.-C., Stappenbeck, T.S., Gaut, J.P., Swamidass, S.J., 2018. Deep learning global glomerulosclerosis in transplant kidney frozen sections. IEEE Trans. Med. Imaging 37 (12), 2718–2728.

Mei, K., Zhu, C., Jiang, L., Liu, J., Qiao, Y., 2020. Cross-stained segmentation from renal biopsy images using multi-level adversarial learning. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 1424–1428.

Park, J., Woo, S., Lee, J.-Y., Kweon, I.S., 2018. Bam: Bottleneck attention module. pp. 1–14, arXiv preprint arXiv:1807.06514.

Polito, M.G., De Moura, L.A.R., Kirsztajn, G.M., 2010. An overview on frequency of renal biopsy diagnosis in Brazil: clinical and pathological patterns based on 9617 native kidney biopsies. Nephrol. Dial. Transplant. 25 (2), 490–496.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.

Seeja, R., Suresh, A., 2019. Deep learning based skin lesion segmentation and classification of melanoma using support vector machine (SVM). Asian Pac. J. Cancer Prev.: APJCP 20 (5), 1555.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. pp. 1–14, arXiv preprint arXiv:1409.1556.

Sornapudi, S., Hagerty, J., Stanley, R.J., Stoecker, W.V., Long, R., Antani, S., Thoma, G., Zuna, R., Frazier, S.R., 2020. EpithNet: Deep regression for epithelium segmentation in cervical histology images. J. Pathol. Inform. 11, 1—10.

Tan, M., Le, Q., 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning. PMLR, pp. 6105–6114.

Weinstein, J.R., Anderson, S., 2010. The aging kidney: physiological changes. Adv. Chronic Kidney Dis. 17 (4), 302–307.

Yu, Z., Feng, C., Liu, M.-Y., Ramalingam, S., 2017. Casenet: Deep category-aware semantic edge detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5964–5973.

Zeng, C., Nan, Y., Xu, F., Lei, Q., Li, F., Chen, T., Liang, S., Hou, X., Lv, B., Liang, D., et al., 2020. Identification of glomerular lesions and intrinsic glomerular cell types in kidney diseases via deep learning. J. Pathol. 252 (1), 53–64.

Zhang, H., Wu, C., Zhang, Z., Zhu, Y., Lin, H., Zhang, Z., Sun, Y., He, T., Mueller, J., Manmatha, R., et al., 2020. Resnest: Split-attention networks. pp. 1–12, arXiv preprint arXiv:2004.08955.

Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J., 2018. Unet++: A nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, pp. 3–11.