

# Segmentation of the Glomerular Region in Pathological Kidney Slides

Débora B. Nascimento<sup>1</sup>, Luciano Oliveira<sup>2</sup>, Washington Santos<sup>3</sup>, Angelo Duarte<sup>4</sup>,  
Kelson R. T. Aires<sup>1</sup>, Rodrigo M. S. Veras<sup>1</sup>

<sup>1</sup>Universidade Federal do Piauí (UFPI) – Teresina, Brasil

<sup>2</sup>Universidade Federal da Bahia (UFBA) – Salvador, Brasil

<sup>3</sup>Centro de Pesquisas Gonçalo Moniz, Fundação Oswaldo Cruz (FIOCRUZ-BA) –  
Salvador, Brasil

<sup>4</sup>Universidade Estadual de Feira de Santana (UEFS) – Feira de Santana, Brasil

{debora.nascimento, kelson, rveras}@ufpi.edu.br  
lrebouca@ufba.br  
washington.santos@fiocruz.br  
angelo.duarte@uefs.br

**Abstract.** *This study proposes a method for segmenting the glomerular region in renal histological images using convolutional neural networks (CNNs) based on U-Net and Sharp U-Net architectures with pretrained backbones. A total of 643 images stained with hematoxylin-eosin (HE), periodic acid-Schiff (PAS), and periodic acid-methenamine silver (PAMS) were evaluated, applying stratified 5-fold cross-validation. The U-Net with VGG-19 achieved the highest mean Dice score (95.45%), followed by the Sharp U-Net with DenseNet201. The results were consistent across staining techniques, with a slight advantage for PAMS and PAS. The method demonstrated accuracy and robustness, highlighting its potential as a diagnostic support tool in nephropathology.*

**Resumo.** *Este estudo propõe um método para segmentação da região glomerular em imagens histológicas renais, utilizando redes neurais convolucionais (CNNs) baseadas nas arquiteturas U-Net e Sharp U-Net com backbones pré-treinados. Foram avaliadas 643 imagens coradas por hematoxilina-eosina (HE), ácido periódico de Schiff (PAS) e ácido periódico-metenamina prata (PAMS), aplicando validação cruzada estratificada em 5 folds. A U-Net com VGG-19 mostrou o maior Dice médio (95,45%), seguida pela Sharp U-Net com DenseNet201. Os resultados foram consistentes entre as colorações, com leve vantagem para PAMS e PAS. O método revelou precisão e robustez, destacando seu potencial como ferramenta de apoio ao diagnóstico em nefropatologia.*

## 1. Introdução

Os rins são órgãos vitais, localizados na parede posterior do abdômen, com a forma de um grão de feijão e o tamanho aproximado de uma mão fechada. Eles exercem funções essenciais ao organismo, incluindo a filtração e eliminação de toxinas do sangue, a regulação da pressão arterial, a manutenção do equilíbrio hidroeletrólítico e acidobásico, além da produção de hormônios que estimulam a formação de células [Guyton and Hall 2017].

Quando esses órgãos deixam de funcionar adequadamente, podem surgir diversas doenças renais, como as nefrolitíases, glomerulopatias, insuficiência renal, nefropatias diabéticas e hipertensivas, entre outras. No Brasil, aproximadamente 21 mil pessoas iniciam anualmente tratamentos por hemodiálise ou diálise peritoneal, porém, são raros os casos de recuperação da função renal, sendo o transplante uma alternativa limitada a poucos pacientes [Sociedade Brasileira De Nefrologia 2023].

Entre as principais causas de disfunção renal estão as glomerulopatias, um grupo de doenças que acometem os glomérulos, estruturas responsáveis pela ultrafiltração do plasma. Essas condições podem ter origem primária ou serem secundárias a outras doenças, como diabetes e doenças autoimunes, como também podem ter natureza aguda ou crônica. Quando não diagnosticadas e tratadas precocemente, podem evoluir para insuficiência renal crônica terminal. Nesses casos, a biópsia renal é frequentemente utilizada pelo nefrologista como uma ferramenta essencial para o diagnóstico preciso e o diagnóstico do tipo de glomerulopatia [Sociedade Brasileira De Nefrologia 2023].

Com os avanços da tecnologia, especialmente no campo da visão computacional, têm sido desenvolvidos sistemas capazes de segmentar e classificar estruturas celulares em imagens médicas [Litjens et al. 2017]. Desse modo, o uso de redes neurais convolucionais (*Convolution Neural Networks* - CNNs) e algoritmos de aprendizado profundo contribuíram para fortalecer a identificação e o estudo histológico de tecidos, fornecendo informações mais precisas e confiáveis para o diagnóstico e o monitoramento de doenças renais aos profissionais da área.

Nesse contexto, esta pesquisa tem como objetivo a segmentação da região glomerular em lâminas de tecido renal, utilizando técnicas de visão computacional e reconhecimento de padrões. O método proposto visa reduzir ao máximo a intervenção manual e a necessidade de ajustes de parâmetros, como também, ser robusto à presença de ruído e às variações de coloração e qualidade das imagens. Dessa forma, espera-se contribuir para a eficiência e precisão no diagnóstico médico das doenças glomerulares.

Este artigo está organizado da seguinte forma: na Seção 2, são apresentados os trabalhos relacionados que contextualizam este estudo; na Seção 3, são descritos os materiais e métodos, incluindo a base de dados, as arquiteturas utilizadas e os procedimentos de avaliação; na Seção 4, são detalhados os principais resultados e análises obtidos; por fim, a Seção 5 apresenta as conclusões e perspectivas futuras, destacando contribuições e potenciais aplicações do método proposto.

## **2. Trabalhos Relacionados**

A utilização de redes neurais profundas, em especial as convolucionais, tem ganhado destaque na análise de imagens médicas, devido à sua capacidade de extrair automaticamente padrões complexos e reduzir a dependência de intervenções manuais [Litjens et al. 2017]. Considerando o avanço contínuo dessa área, esta seção apresenta uma visão geral sobre o desenvolvimento e a consolidação das abordagens baseadas em aprendizado profundo aplicadas à segmentação de imagens biomédicas.

Gadermayr et al. [Gadermayr et al. 2019] propuseram uma abordagem baseada em cascatas de redes neurais convolucionais para a segmentação de glomérulos em lâminas inteiras (WSIs) renais, e obtiveram um índice Dice de 90%, *recall* de 92% e

precisão de 89%, fazendo uso de validação cruzada de tamanho oito. Os autores destacaram que a metodologia é adequada para cenários com objetos esparsos e distribuição de classes altamente desbalanceadas, como ocorre nos WSIs renais.

Altini et al. [Altini et al. 2020] desenvolveram um sistema de segmentação automático com uso das arquiteturas DeepLabv3+ e SegNet. Os modelos atingiram desempenho elevado, com *F-Score* médio superior a 81% e *IoU* ponderado acima de 97%, mostrando-se eficazes na detecção e classificação de glomérulos em imagens coradas por ácido periódico de Schiff (PAS). A abordagem foi apontada como uma ferramenta de suporte à decisão clínica na triagem de órgãos para transplante, com potencial para padronizar e acelerar a análise patológica.

Além disso, Santos et al. [Santos et al. 2021] propuseram uma abordagem híbrida combinando características profundas e texturais para diferenciar glomeruloesclerose e doença de lesões mínimas a partir de imagens histológicas de biópsias renais. O método inclui a extração de descritores de textura (Haralick e geostatísticos) e *deep features* extraídas de CNNs pré-treinadas (VGG-16, VGG-19, Xception e ResNet50), além de uma etapa de seleção de atributos e classificação por *Random Forest*. Em experimentos realizados sobre 104 imagens, a proposta alcançou acurácia de aproximadamente 90% e índice Kappa acima de 80%, demonstrando que a combinação de descritores texturais e profundos apresentou melhor desempenho que cada categoria isoladamente, sobretudo no diagnóstico diferencial automatizado de patologias glomerulares.

Dimitri et al. [Dimitri et al. 2022] propuseram um estudo baseado na arquitetura DeepLabv2 com codificador ResNet101 para a segmentação de glomérulos esclerosados e não esclerosados em imagens histológicas de biópsias renais. O modelo demonstrou boa capacidade de generalização mesmo diante de diferentes colorações histológicas, e destacou-se o uso positivo da coloração CD10 na identificação de glomérulos esclerosados, ajudando a analisar a forma dos tecidos de maneira automática em exames médicos.

Ademais, Santos et al. [Santos et al. 2024] apresentaram um trabalho que utilizou um *ensemble* de cinco redes convolucionais pré-treinadas (VGG-19, Inception-V3, ResNet-50, DenseNet201 e EfficientNetB2) para detecção de glomeruloesclerose em imagens de glomérulos. Nesse trabalho, foram usadas imagens com as colorações hematoxilina-eosina (HE), ácido periódico-metenamina prata (PAMS) e PAS. Cada CNN passou por um ajuste em duas etapas (*shallow* e *deep fine-tuning*) e foram avaliadas 25 combinações de arquiteturas. A estratégia de votação majoritária entre os melhores modelos resultou em desempenho superior aos classificadores individuais, atingindo acurácia de 99% e índice Kappa de 98%. Esses resultados evidenciam o potencial dos *ensembles* de CNNs para promover maior robustez e estabilidade na detecção automatizada de alterações histopatológicas renais.

Oliveira et al. [Oliveira et al. 2025] investigaram diferentes variantes da U-Net aplicadas à segmentação de imagens histológicas em cenários com poucos dados anotados. As arquiteturas U-Net, U-Net++, Sharp U-Net e TransUNet foram comparadas em conjuntos públicos de tecidos tumorais e glandulares. Os resultados mostraram melhor desempenho da U-Net++ e da Sharp U-Net, devido às conexões densas e filtros de realce, enquanto o TransUNet teve métricas inferiores pela maior complexidade e demanda por dados. O estudo também destacou a importância do aumento de dados geométricos para

melhorar a precisão das segmentações.

Diante desse cenário, é possível observar que os métodos baseados em redes neurais têm avançado significativamente na segmentação glomerular, tanto em qualidade de segmentação quanto em aplicabilidade clínica. No entanto, a área ainda enfrenta desafios importantes, como a limitação de bases de dados anotadas, especialmente no contexto de imagens médicas, o que dificulta a generalização dos modelos para diferentes condições, colorações e variações morfológicas. Esses aspectos reforçam a relevância desta pesquisa nessa área de estudo, buscando contribuir para o aprimoramento das técnicas existentes e para o desenvolvimento de métodos mais eficazes e automatizados de segmentação em imagens histológicas renais.

### 3. Metodologia

Esta seção descreve a metodologia adotada nesta pesquisa, que teve como objetivo a segmentação da região glomerular em imagens histológicas renais. São apresentados os detalhes da base de dados utilizada, composta por imagens obtidas a partir de lâminas com diferentes técnicas de coloração, e as arquiteturas escolhidas para o estudo, com ênfase na U-Net e em variações amplamente empregadas na segmentação biomédica, como U-Net++, Attention U-Net e Sharp U-Net.

Além disso, esta seção aborda a incorporação de diferentes *backbones* convolucionais previamente treinados, incluindo EfficientNetB7, DenseNet201, VGG-16, VGG-19 e MobileNetV2. São descritos também os procedimentos de treinamento com validação cruzada estratificada, implementados com as bibliotecas TensorFlow e Keras, assim como os métodos de avaliação quantitativa aplicados para mensurar o desempenho dos modelos.

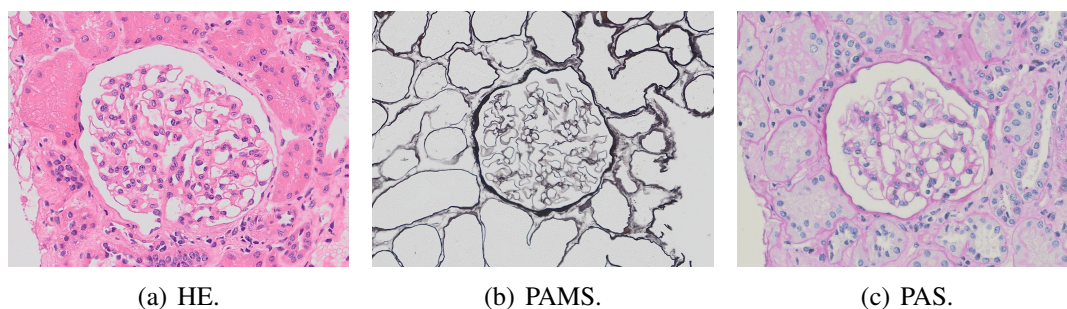
#### 3.1. Base de Imagens

A base de dados utilizada neste trabalho foi fornecida pelo Instituto Oswaldo Cruz (Fiocruz) e contém 45 imagens digitais de fragmentos histológicos de tecido renal obtidas por três técnicas de coloração distintas: HE, PAMS e PAS. Cada técnica de coloração evidencia diferentes aspectos morfológicos do tecido, contribuindo para uma representação visual mais abrangente das estruturas glomerulares.

A Figura 1 apresenta exemplos representativos dessas imagens, destacando as variações de contraste e tonalidade entre as técnicas histológicas. Desse modo, a partir das lâminas originais, foram recortadas automaticamente 643 imagens com resolução de  $1024 \times 768$  pixels, com o objetivo de centralizar os glomérulos no campo da imagem e formar um conjunto experimental adequado ao treinamento e à avaliação dos métodos de segmentação. Essas imagens foram distribuídas da seguinte forma: 219 amostras da coloração HE, 213 da PAMS e 211 da PAS.

Apesar da tentativa de manter todas as estruturas-alvo centralizadas nos recortes, parte das imagens apresentaram glomérulos posicionados próximos às margens. Isto ocorreu devido ao fato de alguns glomérulos estarem localizados na extremidade da lâmina original, o que limita o reposicionamento completo durante o pré-processamento. Essa característica foi mantida no conjunto experimental por refletir condições reais de aquisição e variação na disposição anatômica dos tecidos.





**Figura 1. Imagens do tecido renal centralizadas na região glomerular, contemplando as três técnicas de coloração incluídas no estudo.**

### 3.2. Arquiteturas U-Net e Suas Variações

Entre as principais arquiteturas voltadas à segmentação semântica de imagens biomédicas, a U-Net se destaca como uma solução eficaz e amplamente utilizada. Sua estrutura simétrica e em forma de “U” é composta por dois caminhos principais: um de codificação (*encoder*), formado por camadas de convolução e *pooling*, e outro de decodificação (*decoder*), que recupera a resolução original. Essa abordagem provou ser altamente eficiente mesmo com conjuntos de dados limitados, sendo adotada em aplicações clínicas e científicas [Ronneberger et al. 2015].

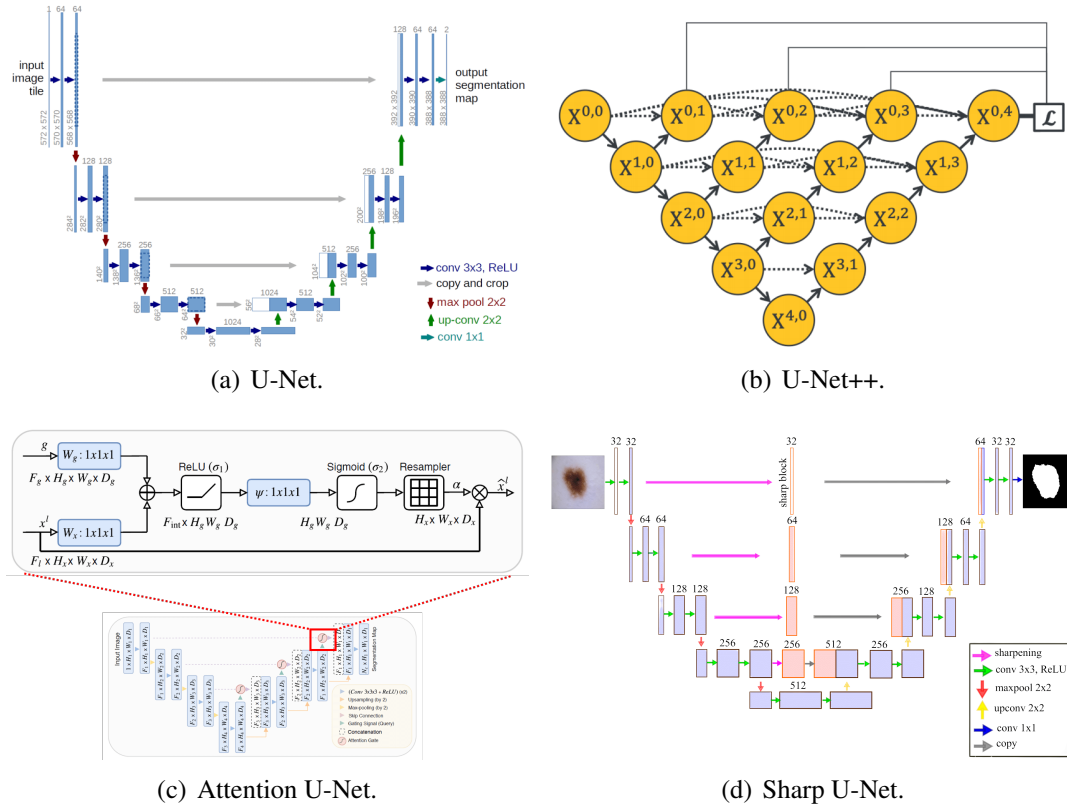
Após sua introdução, surgiram diversas variações. Entre elas, a U-Net++ que preserva a estrutura geral, mas modifica as conexões de atalho (*skip connections*) utilizando conexões densas e aninhadas, o que melhora a fusão de informações e a precisão na segmentação. Além disso, adota supervisão profunda, permitindo treinar redes de diferentes profundidades em conjunto [Zhou et al. 2018, Zhou et al. 2020].

Ademais, a Attention U-Net insere módulos de atenção (*Attention Gates*) nas conexões de atalho, destacando regiões relevantes e reduzindo ativações de fundo [Oktay et al. 2018]. Por fim, a Sharp U-Net é uma variação que aplica uma convolução *depthwise* com um filtro de realce (*sharpening kernel*) antes da fusão dos mapas de características, gerando representações mais nítidas e homogêneas sem aumentar o número de parâmetros treináveis [Zunair and Ben Hamza 2021].

A Figura 2 apresenta exemplos das arquiteturas U-Net, U-Net++, Attention U-Net e Sharp U-Net, retirados dos trabalhos que propuseram cada modelo. As ilustrações mostram como essas variações mantêm a estrutura em “U” e diferem principalmente nas conexões, nos blocos de atenção e nos filtros de realce, que visam aprimorar a segmentação em imagens médicas.

### 3.3. Backbones

Em redes neurais convolucionais para visão computacional, o termo *backbone* refere-se ao modelo de extração de características utilizado como base no *encoder*. Esses modelos são geralmente redes convolucionais profundas pré-treinadas em grandes bases de dados, como o ImageNet. Ao combinar essas redes como *encoders* em arquiteturas como a U-Net e Sharp U-Net, é possível aplicar o conceito de transferência de aprendizado, em que os pesos previamente treinados são reutilizados e, opcionalmente, ajustados para um outro campo, sendo uma estratégia eficaz e recomendada para aplicações médicas,



**Figura 2. Arquiteturas baseadas na U-Net utilizadas na pesquisa: (a) U-Net, (b) U-Net++, (c) Attention U-Net e (d) Sharp U-Net.**

especialmente quando os recursos computacionais e a quantidade de dados anotados são restritos [Tajbakhsh et al. 2016].

Nesta pesquisa, foi utilizada a U-Net e a Sharp U-Net como arquiteturas base, sendo modificadas para incluir diferentes *backbones*, sendo estes, EfficientNetB7, DenseNet201, VGG-16, VGG-19 e MobileNetV2. A Tabela 1 mostra as características das arquiteturas selecionadas como *backbone* para essa pesquisa.

**Tabela 1. Características das arquiteturas utilizadas como *backbone*.**

<i>Backbone</i>	Profundidade	Parâmetros (em milhões)	Acurácia Top-5
MobileNetV2	105	3,5	90,1%
VGG-16	16	138,4	90,1%
VGG-19	19	143,7	90,0%
DenseNet201	402	20,2	93,6%
EfficientNetB7	438	66,7	97,0%

### 3.4. Validação Cruzada Estratificada

Para estimar de forma mais confiável o desempenho e a capacidade de generalização dos modelos, foi adotada a estratégia de validação cruzada estratificada *k-fold*. Nessa abordagem, o conjunto de dados é particionado em *k* subconjuntos de tamanho aproximadamente igual, assegurando que cada *fold* preserve proporcionalmente a distribuição das

colorações presentes na base original. Em cada iteração, um dos *folds* é utilizado como conjunto de teste, enquanto os demais são destinados ao treinamento e à validação. Ao final, os resultados são obtidos pela média das métricas calculadas em todas as execuções. Essa estratégia contribui para reduzir a influência de divisões aleatórias dos dados e torna a comparação entre os modelos mais robusta e confiável [Kohavi 1995].

Neste trabalho, adotou-se  $k=5$ , com a divisão das imagens na proporção de 72% para treinamento, 8% para avaliação e 20% para teste. As imagens originais, com resolução de  $1024 \times 768$  pixels, foram redimensionadas para  $224 \times 224$  pixels, de forma a atender às restrições dimensionais impostas pelos *backbones* utilizados. O treinamento foi conduzido ao longo de 100 épocas, com *batch size* igual a 2, utilizando o recurso de *Early Stopping* para mitigar o risco de *overfitting*. A implementação foi realizada com as bibliotecas TensorFlow 2.10 e Keras 2.10, no ambiente Python 3.10.16.

Como etapa preliminar deste estudo, a validação cruzada estratificada também foi empregada para comparar diferentes variações da U-Net, com o objetivo de identificar aquelas com maior potencial para segmentação glomerular. Nessa avaliação inicial, a U-Net convencional e a Sharp U-Net apresentaram os maiores valores médios de Dice, motivo pelo qual foram selecionadas para os experimentos principais, nos quais foram combinadas com diferentes *backbones* pré-treinados e submetidas a análises quantitativas mais detalhadas.

### 3.5. Métricas de Avaliação

A avaliação dos modelos de segmentação foi conduzida com base em quatro métricas quantitativas comumente empregadas em tarefas de segmentação semântica: índice de Dice, sensibilidade, especificidade e acurácia. Todas as métricas foram calculadas a partir dos resultados obtidos no conjunto de testes, ao término do treinamento de cada modelo.

O índice de Dice, calcula a similaridade entre a segmentação prevista pelo modelo e a segmentação de referência, sendo particularmente adequado para cenários como imagens histológicas renais. A sensibilidade expressa a proporção de regiões glomerulares corretamente identificadas, ou seja, verdadeiros positivos, enquanto a especificidade indica a proporção de regiões corretamente classificadas como não-glomerulares, isto é, verdadeiros negativos. Por fim, a acurácia corresponde à proporção global de atribuições corretas, considerando todas as regiões da imagem.

## 4. Resultados e Discussão

Nesta etapa, os experimentos conduzidos foram realizados com o objetivo de quantificar o desempenho dos modelos de segmentação aplicados às imagens histológicas renais. As avaliações consideraram diferentes combinações de arquiteturas e *backbones* convolucionais, empregando validação cruzada estratificada para assegurar resultados consistentes e minimizar a influência de variações na divisão dos dados.

Os resultados médios de desempenho de cada configuração, considerando as métricas Dice, sensibilidade, especificidade e acurácia, estão apresentados na Tabela 2. Assim, é possível observar que, entre as combinações testadas, a U-Net com o *backbone* VGG-19 alcançou o maior Dice médio, atingindo  $95,45\% \pm 0,57\%$ , seguida pela Sharp U-Net com DenseNet201 ( $95,23\% \pm 0,68\%$ ). No critério de sensibilidade, também se

**Tabela 2. Resultados médios da validação cruzada estratificada nas arquiteturas U-Net e Sharp U-Net com diferentes *backbones*.**

Arquitetura	Backbone	Dice (%)	Sensibilidade (%)	Especificidade (%)	Acurácia (%)
U-Net	EfficientNetB7	94,99 ± 0,60	93,37 ± 0,64	98,92 ± 0,28	97,65 ± 0,19
	MobileNetV2	90,66 ± 5,54	85,81 ± 9,20	98,95 ± 0,18	95,95 ± 2,12
	VGG-19	<b>95,45 ± 0,57</b>	<b>94,80 ± 0,94</b>	98,83 ± 0,23	<b>97,91 ± 0,15</b>
	VGG-16	95,13 ± 0,77	94,00 ± 0,63	98,89 ± 0,23	97,77 ± 0,21
	DenseNet201	94,96 ± 0,64	93,88 ± 0,71	98,80 ± 0,27	97,68 ± 0,23
Sharp U-Net	EfficientNetB7	94,66 ± 0,60	92,64 ± 0,68	<b>98,98 ± 0,29</b>	97,52 ± 0,19
	MobileNetV2	88,50 ± 5,78	82,28 ± 10,42	98,94 ± 0,44	95,05 ± 2,41
	VGG-19	94,92 ± 0,88	94,27 ± 1,13	98,66 ± 0,21	97,67 ± 0,28
	VGG-16	94,78 ± 0,69	94,10 ± 0,78	98,59 ± 0,17	97,57 ± 0,17
	DenseNet201	95,23 ± 0,68	94,35 ± 0,48	98,84 ± 0,28	97,81 ± 0,21

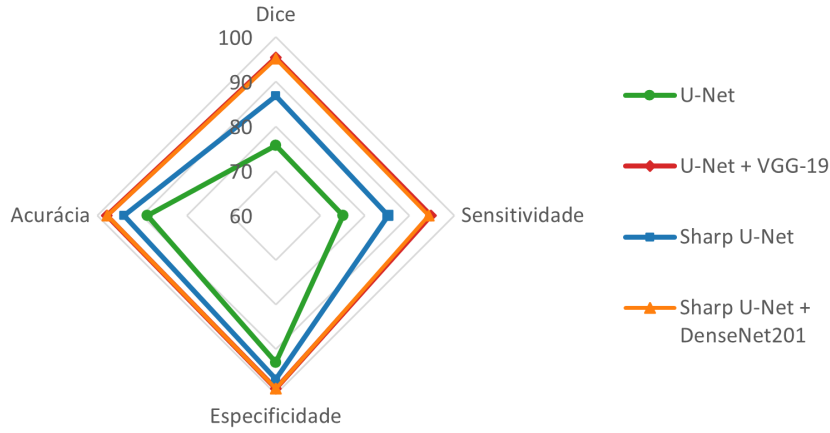
destacaram essas duas configurações, evidenciando elevada capacidade de identificar corretamente as regiões glomerulares. No conjunto dos resultados, todas as variantes apresentaram especificidade superior a 98%, confirmando bom equilíbrio entre detecção de áreas de interesse e redução de falsos positivos, enquanto a acurácia manteve-se elevada e com baixa variação entre os modelos.

Em relação à especificidade, todos os modelos mantiveram índices superiores a 98%, com destaque para a Sharp U-Net com EfficientNetB7, que atingiu 98,98%. A acurácia geral variou de 95,05% a 97,91%, demonstrando estabilidade no desempenho, mesmo considerando diferentes *backbones* e arquiteturas.

De maneira geral, os resultados apontam que tanto a U-Net quanto a Sharp U-Net foram eficazes na segmentação glomerular, com desempenho muito próximo em várias métricas. A leve superioridade da U-Net associada ao VGG-19 e da Sharp U-Net com DenseNet201 sugere que essas combinações apresentam melhor equilíbrio entre sensibilidade e especificidade, sendo, portanto, mais indicadas como alternativas robustas para aplicação prática em imagens histológicas renais.

A Figura 3 apresenta a comparação visual das métricas médias das quatro arquiteturas selecionadas, contemplando tanto as versões base (U-Net e Sharp U-Net) quanto aquelas combinadas com *backbones* pré-treinados (U-Net com VGG-19 e Sharp U-Net com DenseNet201). Nota-se que as variantes com *backbones* alcançaram desempenho superior em todas as métricas avaliadas, enquanto as arquiteturas sem *backbones* apresentaram resultados comparativamente inferiores, com a U-Net obtendo um desempenho menor que a Sharp U-Net. De modo geral, as diferenças entre as melhores combinações foram pequenas, indicando que ambas, mesmo através de combinações de arquitetura e *backbone* diferentes, se mostraram robustas e eficazes na segmentação.

Na Tabela 3, é apresentado os valores do índice Dice, nas arquiteturas U-Net e Sharp U-Net com *backbones*, separados por coloração, sendo possível observar que as colorações PAMS e PAS resultaram nos maiores índices em todas as arquiteturas avaliadas. Já as imagens coradas por HE apresentaram valores de Dice ligeiramente inferiores, ainda que com variação média restrita a aproximadamente 2 pontos percentuais em relação às demais colorações. Esses resultados sugerem que a técnica de coloração exerce influência moderada sobre a acurácia da segmentação, embora as diferenças não tenham sido estatisticamente significativas.



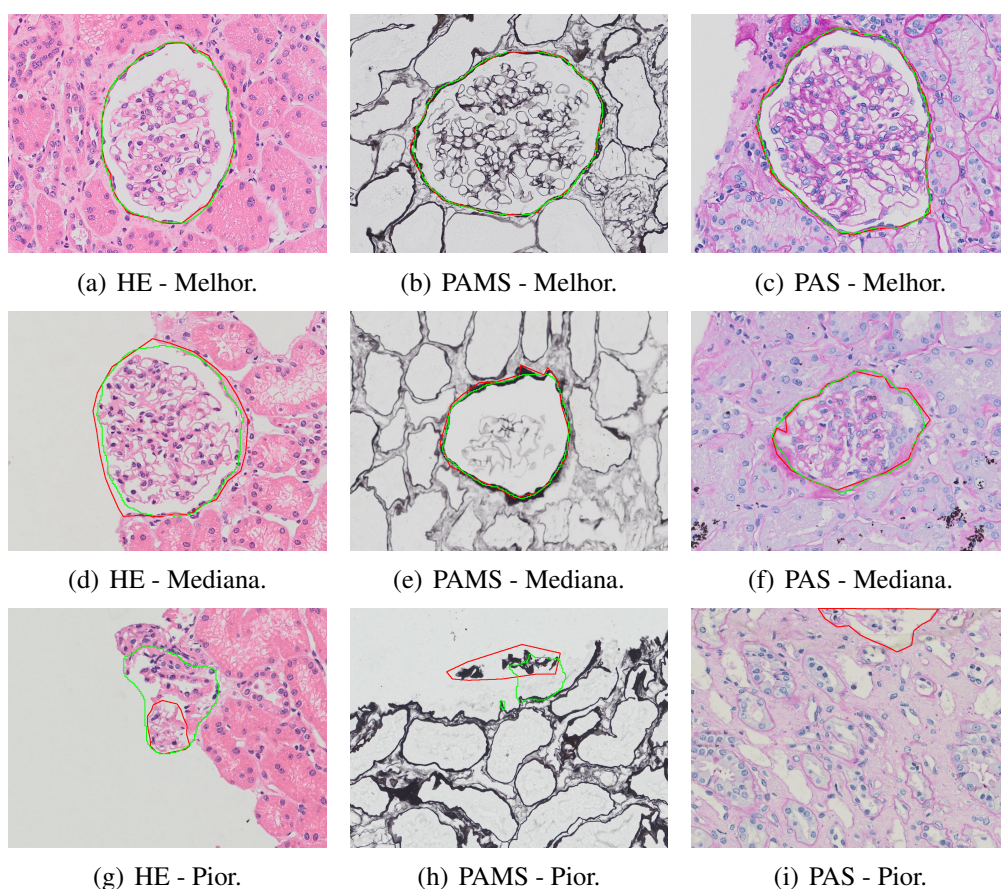
**Figura 3. Comparação visual das métricas médias globais (Dice, sensibilidade, especificidade e acurácia) das arquiteturas U-Net e Sharp U-Net com e sem *backbones* pré-treinados.**

**Tabela 3. Dice médio (%) por coloração histológica (HE, PAMS e PAS) nas arquiteturas U-Net e Sharp U-Net com diferentes *backbones*. Os valores em negrito indicam, para cada combinação de arquitetura e *backbone*, a coloração histológica com maior Dice.**

Arquitetura	Backbone	HE (%)	PAMS (%)	PAS (%)
U-Net	EfficientNetB7	92,34 ± 8,17	93,58 ± 9,12	<b>94,23 ± 9,01</b>
	MobileNetV2	87,61 ± 13,59	85,77 ± 18,11	<b>89,72 ± 11,88</b>
	VGG-19	92,88 ± 7,30	<b>94,79 ± 7,78</b>	94,42 ± 9,33
	VGG-16	92,42 ± 7,57	94,21 ± 9,51	<b>94,22 ± 8,78</b>
	DenseNet201	91,97 ± 9,64	<b>94,28 ± 8,58</b>	94,09 ± 8,91
Sharp U-Net	EfficientNetB7	92,11 ± 8,16	93,08 ± 9,59	<b>94,21 ± 7,44</b>
	MobileNetV2	84,81 ± 14,60	83,51 ± 19,38	<b>87,66 ± 11,96</b>
	VGG-19	92,01 ± 7,55	<b>93,94 ± 9,54</b>	93,81 ± 9,37
	VGG-16	91,90 ± 7,74	<b>93,73 ± 8,23</b>	93,64 ± 9,44
	DenseNet201	92,57 ± 8,13	<b>94,55 ± 8,69</b>	94,41 ± 7,44

De forma complementar, a Figura 4 ilustra exemplos qualitativos das segmentações geradas pela U-Net com VGG-19, contemplando casos representativos das três colorações analisadas. Observa-se que, mesmo em imagens provenientes de diferentes técnicas histológicas, o modelo foi capaz de produzir contornos bem delineados na maioria dos casos, com boa correspondência entre as regiões segmentadas e as máscaras de referência. No entanto, é possível identificar variações de desempenho, especialmente em amostras com menor contraste, com maior heterogeneidade morfológica ou glomérulos situados nas extremidades da imagem, que resultaram em casos com Dice inferior. Essa análise visual corrobora os resultados quantitativos e evidencia a robustez do modelo frente às variações de coloração, embora ainda existam limitações pontuais que podem impactar a qualidade final da segmentação.





**Figura 4. Exemplos qualitativos de segmentações pela U-Net com VGG-19 nas três colorações histológicas. São mostrados casos de melhor, mediano e pior Dice Score em cada coloração, com as regiões segmentadas destacadas em verde e as máscaras de referência em vermelho.**

De forma geral, os experimentos realizados demonstraram que a utilização de *backbones* pré-treinados, em especial a combinação da U-Net com VGG-19 e da Sharp U-Net com DenseNet201, proporcionou ganhos expressivos na precisão da segmentação glomerular. Além disso, os resultados indicaram que a performance dos modelos manteve-se consistente entre diferentes técnicas de coloração histológica, validando a aplicabilidade do método proposto em cenários com heterogeneidade de dados.

## 5. Conclusão

Os experimentos realizados neste estudo mostraram que o uso de *backbones* pré-treinados associados às arquiteturas U-Net e Sharp U-Net contribuiu de forma significativa para melhorar a segmentação da região glomerular em imagens histológicas renais. Entre as configurações avaliadas, a U-Net combinada ao VGG-19 obteve o maior índice médio de Dice (95,45%), seguida de perto pela Sharp U-Net com DenseNet201 (95,23%). Esses resultados evidenciam que diferentes estratégias arquiteturais podem alcançar desempenhos muito próximos quando integradas a técnicas de transferência de aprendizado.

A análise por coloração revelou que as imagens coradas por PAMS e PAS apresentaram ligeira vantagem nos índices de segmentação em comparação à HE. No entanto, a variação entre as colorações foi pequena e não comprometeu a robustez geral dos mode-

los. A validação cruzada estratificada com cinco *folds* reforçou a consistência e a estabilidade dos resultados, indicando boa capacidade de generalização mesmo com um conjunto limitado de dados.

Os exemplos qualitativos demonstraram que os modelos foram capazes de delinear adequadamente os contornos na maioria das amostras, ainda que casos específicos, como imagens com menor contraste ou estruturas próximas das bordas, tenham se revelado como impasse. Dessa forma, destaca-se a importância de futuras investigações voltadas ao aprimoramento da robustez frente a artefatos e variações de qualidade das lâminas.

Como trabalhos futuros, pretende-se explorar estratégias adicionais de pré-processamento, técnicas de aumento de dados e diferentes arquiteturas de CNNs, bem como, realizar o treinamento dos modelos para cada coloração separadamente. Tais avanços podem potencializar ainda mais a acurácia e a robustez frente às variações morfológicas e de coloração. Assim, os achados deste estudo reforçam o potencial das redes profundas como ferramentas de apoio ao diagnóstico e à automação da análise patológica de doenças glomerulares.

## Referências

- Altini, N. et al. (2020). Semantic segmentation framework for glomeruli detection and classification in kidney histological sections. *Electronics*, 9(3):503.
- Dimitri, G. M. et al. (2022). Deep learning approaches for the segmentation of glomeruli in kidney histopathological images. *Mathematics*, 10(11):1934.
- Gadermayr, M. et al. (2019). Cnn cascades for segmenting sparse objects in gigapixel whole slide images. *Comput. Med. Imaging Graph.*, 71:40–48.
- Guyton, A. C. and Hall, J. E. (2017). *Tratado de fisiologia médica*. Elsevier, Rio de Janeiro, 13 edition.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, volume 2, pages 1137–1143. Morgan Kaufmann Publishers Inc.
- Litjens, G. et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88.
- Oktay, O., Schlemper, J., Le Folgoc, L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., and Rueckert, D. (2018). Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*.
- Oliveira, D., Silva, A., Neves, L., Tosta, T., Martins, A., Faria, P., and Nascimento, M. (2025). Investigação de arquiteturas de aprendizagem profunda para segmentação de lesões em imagens histológicas. In *Anais do XXV Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 617–628, Porto Alegre, RS, Brasil. SBC.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer.
- Santos, J., Silva, R., Oliveira, L., Santos, W., Aldeman, N., Duarte, A., and Veras, R. (2024). Glomerulosclerosis detection with pre-trained cnns ensemble. *Computational Statistics*, 39:561–581.

- Santos, J. D., Veras, R. M. d. S., and Silva, R. R. (2021). A hybrid of deep and textural features to differentiate glomerulosclerosis and minimal change disease from glomerulus biopsy images. *Biomedical Signal Processing and Control*, 70:103020.
- Sociedade Brasileira De Nefrologia (2023). Glomerulopatias. Acesso em: 25 mar. 2025.
- Tajbakhsh, N. et al. (2016). Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Transactions on Medical Imaging*, 35(5):1299–1312.
- Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., and 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*, pages 3–11. Springer.
- Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., and Liang, J. (2020). Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging*, 39(6):1856–1867.
- Zunair, H. and Ben Hamza, A. (2021). Sharp u-net: Depthwise convolutional network for biomedical image segmentation. *Computers in Biology and Medicine*, 136:104699.