

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

LOURENÇO MADRUGA BARBOSA

ANÁLISE DE HISTOPATOLOGIA RENAL USANDO DEEP LEARNING

Curitiba PR

2020

LOURENÇO MADRUGA BARBOSA

ANÁLISE DE HISTOPATOLOGIA RENAL USANDO DEEP LEARNING

ANALYSIS OF KIDNEY HISTOPATHOLOGY USING DEEP LEARNING

Trabalho de pesquisa de mestrado apresentada como requisito para obtenção do título de Mestre em Ciências do curso de Pós-Graduação em Engenharia Elétrica e Informática Industrial/CPGEI - Área de Concentração: Engenharia Biomédica.

Orientador: Prof. Dr. Fábio Kurt Schneider

Coorientador: Prof. Dr. Wilson José da Silva

Curitiba PR

2020



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Esta licença permite download e compartilhamento do trabalho desde que sejam atribuídos créditos ao(s) autor(es), sem a possibilidade de alterá-lo ou utilizá-lo para fins comerciais. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Curitiba



LOURENCO MADRUGA BARBOSA

ANÁLISE DE HISTOPATOLOGIA RENAL USANDO DEEP LEARNING

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Ciências da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Engenharia Biomédica.

Data de aprovação: 17 de Dezembro de 2020

Prof Fabio Kurt Schneider, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Dalcimar Casanova, - Universidade Tecnológica Federal do Paraná

Prof.a Karin Braun Prado, Doutorado - Universidade Federal do Paraná (Ufpr)

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 24/02/2022.

Dedico este trabalho inicialmente a Deus e Nossa Senhora Aparecida por sempre estarem ao meu lado, por colocar em meu caminho pessoas excepcionais, que me auxiliaram, direta ou indiretamente no desenvolvimento desta pesquisa, por terem me dado discernimento nas minhas escolhas para vencer todos os obstáculos e desafios desta longa caminhada. Dedico este trabalho também a minha esposa Karla e meus filhos João Gabriel e Flávia por sempre estarem ao meu lado, por terem tido paciência e compreensão nos momentos em que estive distante. Aos meus pais Josselia e Orocil

AGRADECIMENTOS

Agradeço Imensamente todas as pessoas que estiveram ao meu lado e presenciaram esta etapa da minha vida, talvez para alguns isso possa ser apenas um passo, mas para mim esta vitória representa muito na minha vida, o mestrado para mim representava uma realidade distante, mas com a ajuda de muitas pessoas envolvidas o dia a dia me fez constatar que eu estava equivocado. Agradeço a Deus e a Nossa Senhora Aparecida, por ter colocado no meu caminho pessoas maravilhosas, ter me dado discernimento nesta trajetória, que possuiu muitos altos e baixos, e nunca me deixou abalar e desistir. Agradeço a FH, meus gerentes Diogo Coimbra e Sérgio Oliveira, por permitirem um horário flexível de trabalho, que possibilitou a conclusão das matérias para obtenção dos créditos necessários para cumprimento das regras determinadas pelo Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial CPGEI.

Agradeço a todos os discentes e docentes da UTFPR que participaram diretamente ou indiretamente para a finalização desta façanha:

Prof. Fabio Kurt Schneider, por ter confiado no meu potencial desde o momento da nossa entrevista, ter dedicado parte do seu tempo incentivado e ajudando a dar um caminho sensato para o término deste trabalho, com muito respeito e serenidade.

Hamilton e Alexssandro, que viabilizaram esta pesquisa, cedendo parte dos seus conhecimentos referentes a Deep Learning da pesquisa de detecção de levedura que estava sendo efetuada, para que pudéssemos utilizar na detecção de glomérulos.

Todos os professores que dedicaram tempo em sala de aula para nos ensinar e enriquecer e aprimoramento nosso conhecimento, sempre com dedicação e entusiasmo, Prof. Dr. Hugo Vieira Neto, Prof. Dr. Pedro Miguel Gewehr, Prof. Dr. Paulo Jose Abatti, Prof. Dr. Joaquim Miguel Maia, Prof. Dr. Paulo César Stadzisz e ao Prof. Dr. Fabio Kurt Schneider

A todos os amigos, que sempre estiveram do meu lado, dispostos a ajudar, e que nos piores momentos, momentos que pensei até em desistir, não me deixaram abater, a todos os envolvidos, só tenho uma palavra, gratidão.

RESUMO

BARBOSA, Lourenço Madruga. Análise De Histopatologia Renal Usando Deep Learning. 2020. 108 f. Dissertação – Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

Neste trabalho utilizamos redes neurais convolucionais e aprendizagem profunda para detectar objetos de interesse biomédico (i.e., glomérulos) em lâminas inteiras digitalizadas de estudos de patologia renal. A análise de glomérulos é importante para estudos de nefropatias causadas por diabetes, lúpus, uso de drogas, excesso de fármacos, entre outras causas. A digitalização de lâminas inteiras permite a aplicação de uma variedade de técnicas de processamento digital de imagem e uso de inteligência artificial para auxiliar os médicos em diagnósticos. Efetuamos o treinamento de uma rede neural convolucional YOLOv3 de arquitetura com 53 camadas. O treinamento e análise de desempenho da rede contou com 6 experimentos. O conjunto de dados de treinamento, continha um total de 16 lâminas inteiras, divididas em subimagens de 2048x2048, que resultaram em um total de 815 imagens, contendo um total de 2325 glomérulos anotados, para o treinamento, validação e avaliação de desempenho. Para a análise de desempenho da rede treinada (i.e., detecção), foram selecionadas 7 lâminas desconhecidas da rede. Com base no resultado encontrado, é possível afirmar que o resultado deste trabalho é superior aos encontrados na literatura. Acurácia, Precisão, Sensibilidade, Especificidade e F1s de 99,40%, 97,31%, 96,17%, 99,73% e 94,24%, respectivamente, foram obtidas neste trabalho mostrando o potencial para auxiliar nos diagnósticos de exames histopatológicos.

Palavras-chave: Aprendizagem profunda; Redes Neurais; Classificação de glomérulos; Detecção de Glomérulo; Patologia Digital; Processamento digital de imagem.

ABSTRACT

BARBOSA, Lourenço Madruga. Analysis of Kidney Histopathology Using Deep Learning. 2020. 108 f. Dissertation - Graduate Program in Electrical Engineering and Industrial Informatics, Federal Technological University of Paraná. Curitiba, 2020.

In this work we use convolutional neural network and deep learning to detect objects of biomedical interest (i.e., glomeruli) in whole digitalized slides from studies of renal pathology. The analysis of glomeruli is important for studies of nephropathies caused by diabetes, lupus, drug use, excess of drugs, among other causes. The digitization of entire slides allows the application of a variety of digital image processing techniques and the use of artificial intelligence to assist doctors in diagnoses. We carried out the of a YOLOv3 convolutional neural network of architecture with 53 layers. The training and performance analysis of the network had 6 experiments. The training data set contained a total of 16 entire slides, divided into sub-images of 2048x2048, which resulted in a total of 815 images, containing a total of 2325 annotated glomeruli, for training, validation and performance evaluation. For the performance analysis of the trained network (i.e., detection), 7 unknown slides of the network were selected. Based on the result found, it is possible to state that the result of this work is superior to those found in the literature. Accuracy, Precision, Sensitivity, Specificity and F1s of 99.40%, 97.31%, 96.17%, 99.73% and 94.24%, respectively, were obtained in this work showing the potential to assist in the diagnosis of histopathological exams.

Keywords — Glomerulus Classification; Glomerulus Detection; Deep Learning, Neural Networks; Digital Pathology; Image processing.

LISTA DE ILUSTRAÇÕES

Figura 1 - Ilustração das estruturas internas do rins.....	19
Figura 2 - Estrutura interna de um néfron.....	20
Figura 3 - Representação esquemática corpúsculo renal	21
Fotografia 1 - Representação fotográfica de fixação de tecido.	23
Fotografia 2 - Representação fotográfica de desidratação da peça.	24
Fotografia 3 - Representação fotográfica da Diafanização.....	24
Fotografia 4 - Representação fotográfica do emblocamento do tecido	25
Fotografia 5 - Representação fotográfica do posicionamento de um bloco.....	26
Fotografia 6 - Representação fotográfica de um seccionamento de um bloco.	26
Fotografia 7 - Representação fotográfica da separação e pesca.....	27
Fotografia 8 - Representação fotográfica do processo de distensão e pesca ..	27
Fotografia 9 - Representação fotográfica da desparafinação	28
Fotografia 10 - Representação fotográfica da hidratação e lavagem	29
Fotografia 11 - Representação fotográfica do banho de hematoxilina	30
Fotografia 12 - Representação fotográfica do banho de eosina.....	31
Fotografia 13 - Representação fotográfica dos banhos em solução de xilol.....	31
Fotografia 14 - Representação fotográfica dos materiais utilizados	32
Fotografia 15 - Representação fotográfica da limpeza secagem e etiquetação.....	32
Figura 4 - Tela de exemplo do software Aperio ImageScope	34
Figura 5 - Representação em blocos da lâmina montada.....	34
Figura 6 - Representação de uma imagem com resolução de 83 x 55	35
Figura 7 - Representação de uma imagem em preto (0) e branco (255).....	36
Figura 8 - Representação de uma imagem em escala de cinza.	36
Figura 9 - Representação de como é formada uma imagem digital	37
Figura 10 - Representação em blocos dos passos para processamento	38
Figura 11 - Representação de um neurônio humano.....	42
Figura 12 - Representação da estrutura de um neurônio artificial.	43
Figura 13 - Representação do gráfico gerado com a aplicação das Equações.....	44
Figura 14 - Representação do gráfico gerado com a aplicação das equações ..	45
Figura 15 - Representação do gráfico gerado com a aplicação das equações ..	46
Figura 16 - Representação do gráfico gerado com a aplicação das equações ..	47
Figura 17 - Representação gráfica do aumento de popularidade da pesquisa.	47
Figura 18 - Representação gráfica de rede neural simples e de múltiplas camadas.....	48
Figura 19 - Representação do sistema de funcionamento YOLO.....	50
Figura 20 - Representação de tensores unidimensional e bidimensional	51
Figura 21 - Representação de aplicação de filtro 3x3 em uma imagem 6x6.	51
Figura 22 - Representação das caixas delimitadoras	52
Figura 23 - Representação da distribuição das camadas na rede YOLOv3.....	53
Figura 24 - Fragmento da imagem com verdadeiro positivo.....	55
Figura 25 - Fragmento de imagem com um falso positivo.	56
Figura 26 - Fragmento de imagem com falso negativo.....	56
Figura 27 - Representação das áreas calculadas.....	57
Figura 28 - Representação do download da imagem Case10_G.....	61
Figura 29 - Representação da imagem Case10_G após o recorte.....	62
Figura 30 - Exemplificação dos 5.168 fragmentos de imagens Bio-Atlas.	63
Figura 31 - Representação das 60 imagens de 2048x2048.....	63

Figura 32 - Representação da imagem anotada através do software YBAT.....	64
Figura 33 - Representação da imagem e do arquivo gerado pelo programa YBAT	65
Figura 34 - Fragmentos das lâminas utilizadas para a avaliação de desempenho.	70
Figura 35 - Fragmento da lâmina Imagem_1307.....	72
Figura 36 - Fragmento da lâmina Imagem_1111.....	73
Gráfico 1 - Demonstração dos resultados dos experimentos.....	76
Gráfico 2 - Comparação entre as métricas de todos os experimentos	78
Figura 37 - Imagem caso 8 utilizada para treinamento	86
Figura 38 - Imagem caso 10 utilizada para o treinamento da rede	86
Figura 39 - Imagem caso 11 utilizada para o treinamento da rede	87
Figura 40 - Imagem caso 12 utilizada para o treinamento da rede	87
Figura 41 - Imagem caso 13 utilizada para o treinamento da rede	88
Figura 42 - Imagem caso 14 utilizada para o treinamento da rede	88
Figura 43 - Imagem caso 15 utilizada para o treinamento da rede	89
Figura 44 - Imagem 837 imagem de biópsia renal utilizada para treinamento... 91	
Figura 45 - Imagem 1060 imagem de biópsia renal utilizada para treinamento. 91	
Figura 46 - Imagem 1267 imagem de biópsia renal utilizada para treinamento. 92	
Figura 47 - Imagem 1292 imagem de biópsia renal utilizada para treinamento. 92	
Figura 48 - Imagem 1298 imagem de biópsia renal utilizada para validação..... 93	
Figura 49 - Imagem 569 imagem de biópsia renal utilizada para avaliação. 93	
Figura 50 - Imagem 1307 imagem de biópsia renal utilizada para avaliação. 94	
Figura 51 - Imagem 603 imagem de biópsia renal utilizada para avaliação 94	
Figura 52 - Imagem 603 imagem de biópsia renal utilizada para avaliação 95	
Figura 53 - Imagem 849 imagem de biópsia renal utilizada para avaliação 95	
Figura 54 - Apresenta a lâmina completa 1057 utilizada para o treinamento 97	
Figura 55 - Apresenta a lâmina completa 33814 utilizada para o treinamento .. 97	
Figura 56 - Apresenta a lâmina completa 134159 utilizada para o treinamento. 98	
Figura 57 - Apresenta a lâmina completa 149335 utilizada para o treinamento 98	
Figura 58 - Apresenta a lâmina completa 1111 utilizada para a avaliação..... 99	
Figura 59 - Apresenta a lâmina completa 33836 utilizada para a avaliação 99	
Figura 60 - Apresenta a lâmina completa 43267 utilizada para a avaliação..... 100	
Figura 61 - Site Rosai's para download da imagem. 105	

LISTA DE TABELAS

Tabela 1 - Hiperparâmetros configurados na para o treinamento, validação e análise de desempenho.....	54
Tabela 2 - Demonstra as lâminas utilizadas, os repositórios utilizados para download das imagens número total de imagens de 2048x2048 e o total de glomérulos anotados por imagens.....	65
Tabela 3 - Tabela com informações da base de imagens utilizadas para o treinamento e validação dos experimentos A e E.....	66
Tabela 4 - Tabela com informações da base de imagens utilizadas para o treinamento e validação do experimento B.....	67
Tabela 5 - Tabela com informações da base de imagens utilizadas para o treinamento e validação do experimento C.....	67
Tabela 6 - Tabela com informações da base de imagens utilizadas para o treinamento e validação do experimento D.....	68
Tabela 7 - Tabela com informações da base de imagens utilizadas para o treinamento e validação do experimento F.....	68
Tabela 8 - Imagens do repositório Bio-Atlas utilizadas para avaliação de desempenho da rede.....	69
Tabela 9 - Imagens do repositório Leeds utilizadas para avaliação de desempenho da rede.....	69
Tabela 10 - Resultados obtidos por (GALLEGO et al., 2018).....	71
Tabela 11 - Resultados obtidos por (LO et al., 2018).....	71
Tabela 12 - Apresentação dos valores obtidos para o experimento E para cada lâmina (i.e., 1111, 66836 e 43267) com seus respectivos parâmetros (i.e., Tot. de imagens, Tot. Glomérulos, FP, FN, TP e Tn).....	73
Tabela 13 - Apresentação das métricas calculadas para o experimento E.....	74
Tabela 14 - Resultados obtidos nos experimentos de A até F efetuados com o framework YOLO utilizados todas as imagens das Tabelas 9 e 10, as duas últimas colunas são utilizadas para mostrar as informações dos artigos utilizados na comparação.....	74

LISTA DE ABREVIATURAS

AI	<i>Artificial Intelligence</i> (Inteligência Artificial)
ANN	<i>Artificial Neural Network</i> (Rede Neural Artificial)
BD	Banco de dados
BMP	<i>Bit Map Picture</i> (Imagem em Mapa de Bit)
DL	<i>Deep Learning</i> (Aprendizagem Profunda)
GIF	<i>Graphics Interchange Format</i> (Formato para Intercâmbio de Imagem)
H&E	Hematoxilina e Eosina
JPEG	<i>Joint Photographi Experts Group</i> (Grupo Misto de Especialista Fotográficos)
MEC	Matriz extracelular
ML	<i>Machine Learning</i> (Aprendizado de Máquina)
NN	<i>Neural Network</i> (Rede Neural)
PAS	<i>Periodic acid-reactive Schiff</i>
PDI	Processamento digital de imagem
PNG	<i>Portable Network Graphics</i> (Imagens Portáteis para Rede)
RGB	<i>Red, Green e Blue</i> (Vermelho, Verde e Azul)
RNA.	Ácido Ribonucleico
TIF	<i>Tagged Image Format</i> (Formato de Imagem com Rótulo)

LISTA DE SÍMBOLO

A	Acurácia
E	Especificidade
Ef	Eficiência
$F1_s$	<i>F1-Score</i>
P	Pressão
f	Função
P	Precisão
S	Sensibilidade
VPN	Valor Preditivo Negativo
(x, y)	Coordenadas
Σ	Somatória
TN	Verdadeiro negativo (<i>True negative</i>)
TP	Verdadeiro positivo (<i>True positive</i>)
FP	Falso positivo (<i>True positive</i>)
FN	Falso negativo (<i>False negative</i>)
H	Altura (<i>Height</i>)
W	Largura (<i>Width</i>)
N_G	Número de glomérulos
\overline{H}	Média da altura
\overline{W}	Média da largura
σ_H	Desvio padrão da altura
σ_W	Desvio padrão da largura
N_N	Número de neurônios

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Objetivos	16
1.1.1	Objetivos Específicos	17
1.1.2	Estrutura da dissertação	17
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	Rins	18
2.1.1	Néfron	19
2.1.2	Corpúsculo renal	20
2.2	Histopatologia tradicional e digital	22
2.2.1	Histopatologia tradicional	22
2.2.2	Fixação	22
2.2.3	Desidratação	23
2.2.4	Diafanização	24
2.2.5	Inclusão ou Emblocamento	25
2.2.6	Posicionamento e Microtomia	25
2.2.7	Separação e pesca dos cortes	26
2.2.8	Distensão e pesca dos cortes	27
2.2.9	Desparafinação	27
2.2.10	Hidratação e Lavagem	28
2.2.11	Coloração	29
2.2.12	Montagem, limpeza, secagem e identificação	31
2.3	Histopatologia digital	33
2.4	Imagem	35
2.4.1	Cor em imagem digital	35
2.4.2	Processamento digital de imagem	37
<u>2.4.2.1</u>	<u>Aquisição da imagem</u>	<u>38</u>
<u>2.4.2.2</u>	<u>Digitalização da imagem</u>	<u>38</u>
<u>2.4.2.3</u>	<u>Segmentação</u>	<u>38</u>
<u>2.4.2.4</u>	<u>Extração de atributos</u>	<u>38</u>
<u>2.4.2.5</u>	<u>Classificação e reconhecimento</u>	<u>39</u>
2.4.3	Processo humano de identificação	39
<u>2.4.3.1</u>	<u>Reconhecimento</u>	<u>39</u>
<u>2.4.3.2</u>	<u>Processamento de imagem de um olho humano</u>	<u>40</u>

2.5	Inteligência artificial	40
2.5.1	Aprendizado de máquina.....	41
2.5.2	Redes neurais artificiais.....	42
<u>2.5.2.1</u>	<u>Modelo de um neurônio humano</u>	<u>42</u>
<u>2.5.2.2</u>	<u>Modelo de um neurônio artificial</u>	<u>43</u>
<u>2.5.2.3</u>	<u>Funções de ativação</u>	<u>44</u>
2.5.3	Aprendizado profundo.....	47
2.5.4	YOLOv3.....	49
2.6	Métricas para avaliação de desempenho	55
3	MATERIAIS E MÉTODOS	59
3.1	Base de dados	59
3.1.1	Base de dados.....	59
3.1.2	Base Bio-Atlas.....	59
3.1.3	Base University of Leeds.....	60
3.2	Preparo das imagens	61
3.2.1	Anotação das imagens.....	64
3.3	Treinamento	65
3.4	Avaliação de desempenho	69
4	RESULTADOS E DISCUSSÃO	71
5	CONCLUSÕES E TRABALHOS FUTUROS	79
	REFERÊNCIAS	81
	APÊNDICE A - Bases de imagens Juan Rosai's	85
	APÊNDICE B - Bases de imagens Bio-Atlas	90
	APÊNDICE C - Bases de imagens LEEDS	96
	APÊNDICE D - Códigos	101

1 INTRODUÇÃO

A Patologia é um ramo da medicina dedicado ao estudo e análise de tecidos e órgãos, ciência que estuda e busca entender quais são os mecanismos envolvidos nas causas das doenças e quais são as alterações morfológicas e funcionais das células e tecidos doentes (KUMAR, ABBAS, e ASTER, 2013). Inúmeras doenças são causadoras ou mediadoras de nefropatias, lesões ou doenças nos rins. Diabetes, lúpus, uso prolongado ou em excesso de fármacos, uso de drogas de abuso (e.g., heroína, cocaína), hipertensão arterial, tumores, amiloidose e infecções como malária, sífilis, hepatite B e HIV, entre outras doenças que podem causar lesões em tecidos renais. Para estes casos, é comum se efetuar a biópsia (i.e., coleta de amostra de tecido para análise). Para ser analisado, o tecido necessita passar por um processo de fixação, preservação do material contra micro-organismos, desidratação, remoção da substância utilizada na fixação, diafanização, remoção do álcool utilizado para desidratação, embocamento (i.e., o material é colocado na parafina para se tornar um bloco consistente para o corte histológico), laminação histológica em um micrótomo. Estas fatias de tecido são então transferidas para a lâmina de vidro e recebem a coloração (e.g., hematoxilina e eosina em estudos com glomérulos), para a típica análise usando microscopia óptica (i.e., o médico neste tipo de análise fica restrito a uma pequena área do tecido, com base no foco do microscópio) (KUMAR, ABBAS, e ASTER, 2013).

Os avanços tecnológicos têm promovido diversas melhorias na área da medicina, uma delas é a disseminação do uso de scanners para digitalização das lâminas com os tecidos. Esta técnica de digitalização das lâminas de vidro com tecido está sendo muito utilizada na patologia digital (PD) e permite aos médicos efetuarem análises de imagens de lâminas inteiras com suporte computacional (isto é, algoritmos dedicados e alto poder computacional) e monitores de alta definição. Com a digitalização das lâminas é criado um repositório de imagens de PD, que pode ser compartilhado entre médicos, pesquisadores e educadores em todo mundo. Isso possibilita a aplicação de uma variedade de técnicas de processamento digital de imagens (PDI) nas imagens de PD, como classificação e segmentação. O desenvolvimento de softwares tem se mostrado promissor, têm a capacidade de auxiliar nas tarefas mais trabalhosas e tediosas de quantificação servindo também como um registro de análises supervisionada auxiliando os patologistas na tomada de decisão.

A. Janowwczyk e A. Madadhushi, recentemente apresentaram um estudo de caso de uso utilizando aprendizado profundo (i.e., *Deep Learning* DL) para análise de PD. O DL é uma subárea do aprendizado de máquina (i.e., *Machine Learning* ML), representa aprendizagem de ponta a ponta (i.e., E2E), treinamento de sistema de aprendizagem, complexo ou não, representado por um modelo de aprendizado por característica (i.e., *Feature Learning*), conjunto de técnicas que permite um sistema descobrir de forma automática características necessárias para detecção ou classificação através de dados brutos. Normalmente, quando o conjunto de treinamento contém um conjunto adequado de padrões, estas redes tendem a apresentar uma boa generalização no conjunto de teste. DL têm se mostrado uma ferramenta com grande potencial, podendo unificar diversas tarefas produzindo resultados excelentes em áreas como contagens de células ou mitoses, classificação de tecidos, classificação de tumores e coloração imuno-histoquímica (JANOWCZYK e MADABHUSHI, 2017).

Trabalhos recentes apresentam o interesse do uso de processamento de imagens e IA para análise glomerular. M. Temerinac-Ott (2017) realizam a detecção de estruturas glomerulares em imagens digitalizadas de slides inteiros para lâminas histopatológicas com coloração histoquímica e imuno-histoquímica múltipla usando uma rede neural convolucional (CNN). T. Kato (2005) propôs um descritor para a detecção glomerular em imagens renais (i.e., SEGMENTAL HOG). R. Marée (2006) apresentou uma arquitetura utilizando os softwares Icy e Cytomine para detector glomérulos em vários centros de análise histopatológica. G. Litjens (2016) apresentam o uso de aprendizagem profunda como técnica para melhoria da acurácia em diagnósticos histopatológicos. Neste trabalho, pretende-se apresentar um estudo onde arquiteturas de aprendizado profundo possam demonstrar quantitativamente a viabilidade de utilização de uma rede de aprendizado profundo para analisar e identificar objetos de interesse (i.e., glomérulo) em lâminas digitais inteiras de biópsia renal.

1.1 Objetivos

O objetivo deste estudo é demonstrar quantitativamente a viabilidade de utilização de uma rede de aprendizado profundo para analisar e identificar objetos de interesse (i.e., glomérulo) em lâminas digitais inteiras de biópsia renal.

1.1.1 Objetivos Específicos

Os objetivos específicos deste estudo são:

- Identificar estruturas glomerulares em biopsias renais com Deep Learning, utilizando base de dados de imagens de lâminas digitais inteiras. Permitindo desta forma, demonstrar quantitativamente a viabilidade de utilização de uma rede de aprendizado profundo.
- Adaptar a base de dados de imagens e o framework You Only Look Once (YOLO) utilizando Python para o treinamento que permita avaliação quantitativa da identificação automática de glomérulos.
- Avaliar quantitativamente os glomérulos existentes em imagens de lâminas inteiras com arquitetura de aprendizado profundo utilizando como métrica (e.g., precisão, sensibilidade, acurácia, especificidade, eficiência, valor preditivo negativo, F1-score), e comparar com resultados apresentados na literatura.

1.1.2 Estrutura da dissertação

Este documento está organizado nos seguintes em:

- Introdução, apresenta a contextualização deste trabalho, através da introdução e seus objetivos;
- Fundamentação teórica, apresenta a fundamentação e os principais conceitos utilizados para o desenvolvimento desta pesquisa;
- Materiais e métodos, descreve todos os materiais e os métodos aplicado no desenvolvimento da pesquisa;
- Resultados, apresenta os resultados obtidos com os experimentos realizados para avaliação do framework YOLOv3;
- Conclusões, contém as conclusões e propostas de trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta a fundamentação teórica desta pesquisa. Inicialmente apresentamos os rins e a existência de glomérulos em tecidos renais pois este é o objeto de análise desta pesquisa. Na sequência são apresentados fundamentos de processamento de imagens que permitam a automatização de identificação de glomérulos em imagens de lâminas completas de histopatologia digital com uso de *Deep Learning*.

2.1 Rins

Os rins são órgãos complexos que possuem um tamanho típico de aproximadamente 10 a 12 cm de comprimento por 5 a 7 cm de largura com massa de aproximadamente 120 a 170 g cada, tem um formato semelhante a um grão de feijão e estão situados na parede posterior do abdômen, as bordas côncavas dos rins ficam voltadas para a coluna vertebral (TORTORA e NIELSEN, 2013).

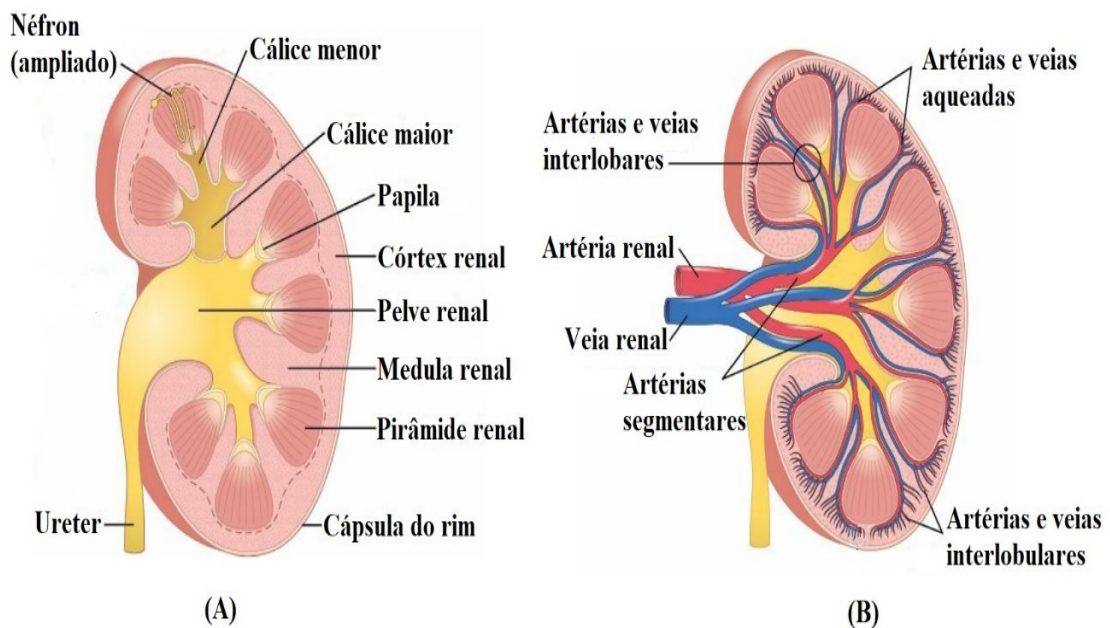
A cavidade onde entra os vasos sanguíneos, vasos linfáticos, ureter e nervos é chamado de hilo. No hilo se encontra o tecido adiposo e os cálices renais que formam a pelve renal, o rim é revestido pela cápsula do rim, um tecido fibroso e resistente que tem como função proteger as estruturas internas, as duas principais regiões dos rins são o córtex renal e a medula renal. A medula renal pode ser dividida entre 8 e 18 pirâmides renais, um tecido em forma de cone, onde a base de cada pirâmide se conecta na papila, que por sua vez está conectada a pelve renal pelo cálice menor que está conectado ao cálice maior, duas divisões da pelve renal que tem a função de coletar e conduzir a urina pelo ureter até a bexiga onde ficara armazenada até a micção. A figura 1 (A) apresenta as estruturas descritas acima (GUYTON e HALL, 2017; TORTORA e NIELSEN, 2013).

O fluxo sanguíneo renal corresponde a aproximadamente 22% do débito cardíaco (volume de sangue sendo bombeado pelo coração em um minuto), que corresponde a 1.100mL/min. A artéria renal entra pelo hilo e se divide formando a artéria segmentar que por sua vez se divide em artérias interlobares, arqueadas, interlobulares (chamadas de artérias radiais) e aferentes, terminando nos capilares glomerulares, onde neste momento começa a filtração para iniciar a formação da urina (GUYTON e HALL, 2017; TORTORA e NIELSEN, 2013).

Neste ponto as extremidades dos capilares glomerulares se juntam a arteríola

eferente formando uma segunda rede de capilares, os peritubulares, fazendo com que o sangue flua para o sistema venoso que formam as veias interlobular, arqueada, interlobar e renal, saindo do rim, pelo hilo junto com o ureter e a artéria renal. A figura 1 (B) ilustra a estrutura interna e externa do fluxo sanguíneo (GUYTON e HALL, 2017; JUNQUEIRA e CARNEIRO, 2013).

Figura 1 – Ilustração das estruturas internas do rim imagem (A) ilustra as estruturas internas e externas com as principais unidades de filtração captação e condução da urina até o rim, a imagens (B) ilustra as estruturas de entrada e saída do fluxo sanguíneo e suas ramificações internas.

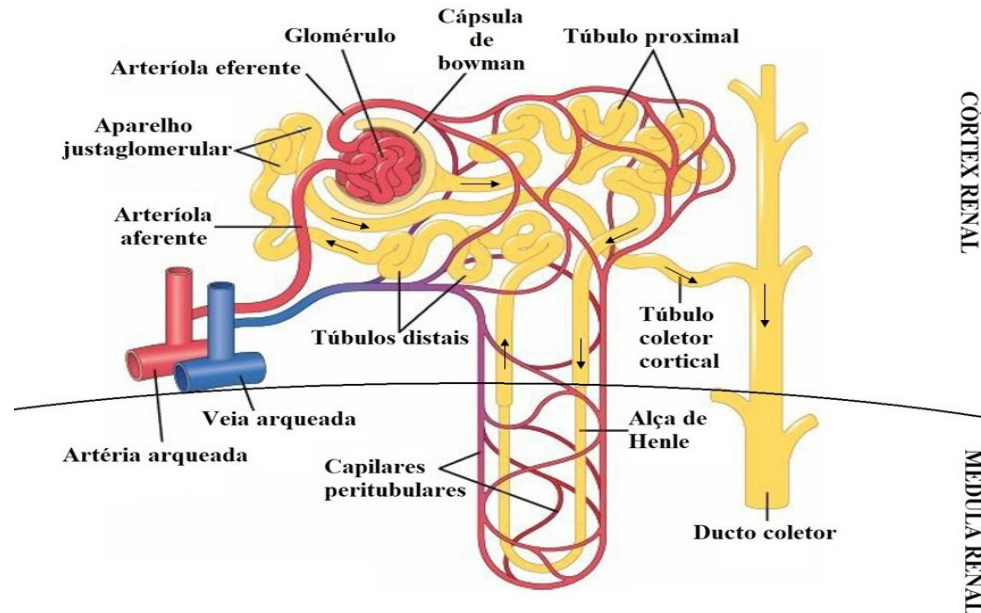


Fonte: Adaptado de GUYTON e HALL (2017, p. 971 e 974)

2.1.1 Nêfron

Nêfrons são as unidades funcionais dos rins, responsável pela filtração do sangue, absorção de água, sais minerais entre outras substâncias, cada rim possui entre 600.000 e 1 milhão de néfrons, cada néfron é composto por várias estruturas, cada uma com uma finalidade. A figura 2 ilustra as estruturas do néfron, (GUYTON e HALL, 2017; JUNQUEIRA e CARNEIRO, 2013).

Figura 2 – Estrutura interna de um néfron.



Fonte: Adaptado de GUYTON e HALL, 2017; JUNQUEIRA e CARNEIRO (2013, pg. 974 e 373)

O corpúsculo renal, em algumas literaturas também chamado de Malpighi, é composto pelo glomérulo, uma rede de capilares ligados a arteríola eferente e aferente, envolto pela capsula de Bowman. Este corpúsculo renal tem a função de filtrar e reter substâncias, o filtrado gerado é direcionado através do túbulo proximal que está ligado a capsula de Bowman (GUYTON e HALL, 2017; JUNQUEIRA e CARNEIRO, 2013). O túbulo proximal, tem a função de reabsorção, é nele que a maior parte da água que foi eliminada é reabsorvida, embora todos os ductos participem desta reabsorção (GUYTON e HALL, 2017; JUNQUEIRA e CARNEIRO, 2013). A alça de Henle, consiste em um ramo descendente e ascendente, ligando o túbulo proximal ao túbulo distal, no ramo descendente parte da água é reabsorvida já no ramo ascendente ocorre a absorção de íons, sódio e cloreto (GUYTON e HALL, 2017; JUNQUEIRA e CARNEIRO, 2013). O túbulo distal tem a função de manter o equilíbrio acidobásico do organismo (GUYTON e HALL, 2017; JUNQUEIRA e CARNEIRO, 2013). Os ductos coletores recebem a urina formada nas estruturas anteriores e se unem a outros ductos formando um ducto maior que finaliza na pelve renal (GUYTON e HALL, 2017; JUNQUEIRA e CARNEIRO, 2013).

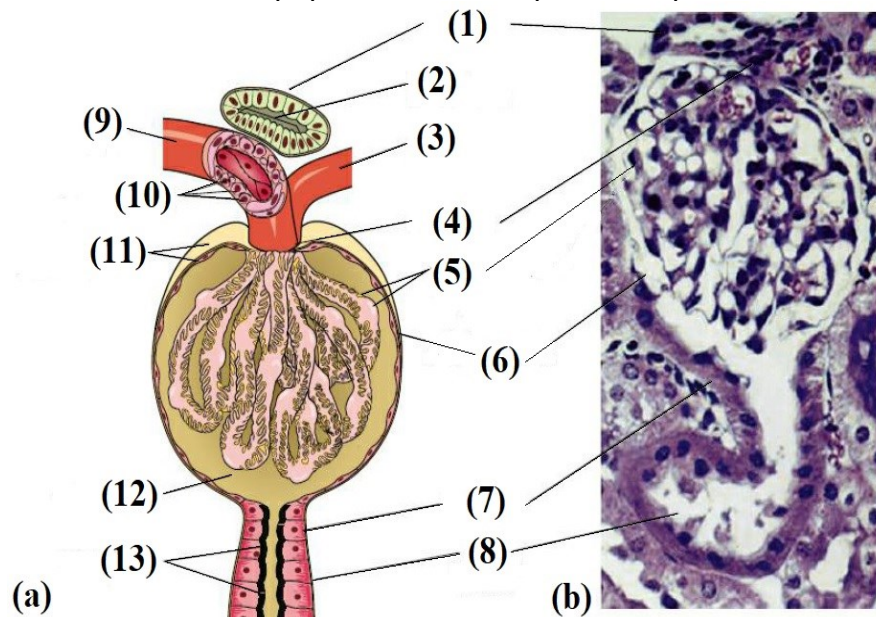
2.1.2 Corpúsculo renal

O corpúsculo renal é uma das principais estruturas dos rins, é nele que grande

quantidade de líquidos são filtrados do sangue, possui cerca de 200 μm de diâmetro, é composto pelo glomérulo, uma rede de capilares também chamada de tufo capilar semelhante a um novelo de lã, envolvida pela capsula de Bowman, constituída pelos folhetos visceral ou interno voltado para o glomérulo e parietal ou externo, formando o limite do corpúsculo renal. O espaço entre os folhetos recebe o nome de espaço capsular, é neste espaço que o líquido filtrado dos capilares flui até o túbulo proximal (GUYTON e HALL, 2017; JUNQUEIRA e CARNEIRO, 2013).

O sangue arterial circula nos capilares, entrando pelo polo vascular através da arteríola aferente e sai pela arteríola eferente, que tem também a função de regular a pressão hidrostática, isso, porque a arteríola eferente possui uma maior quantidade de músculos lisos. A arteríola aferente possui conexão direta com o vaso eferente, o que permite que o sangue passe mesmo sem circular pelo glomérulo. O polo urinário é por onde sai o filtrado glomerular, do início do túbulo contorcido proximal. A figura 3 (a) demonstra a representação esquemática do corpúsculo renal e a (b) a fotomicrografia que nos permite ver detalhes das estruturas de um corpúsculo renal, (GUYTON e HALL, 2017; JUNQUEIRA e CARNEIRO, 2013).

Figura 3 – Representação esquemática corpúsculo renal (a) e fotomicrografia (b), e a descrição das suas partes, (1) Túbulo distal, (2) Mácula densa do túbulo distal, (3) Arteríola eferente, (4) Polo vascular, (5) Folheto visceral da cápsula de Bowman (podócitos), (6) Folheto parietal da cápsula de Bowman, (7) Polo urinário, (8) Túbulo contorcido proximal (9) Arteríola aferente, (10) Células justaglomerulares, (11) folheto parietal da cápsula de Bowman, (12) Espaço capsular e (13) Orla em escova (microvilos).



Fonte: Adaptado de JUNQUEIRA e CARNEIRO (2013, pg. 373)

2.2 Histopatologia tradicional e digital

Histopatologia é o termo cujo sentido tem origem do latim *histo* (i.e., tecido), *phatos* (i.e., doença) e *logos* (i.e., estudo). Esta ciência é um dos ramos da medicina, dedicada ao estudo e análise das causas e alterações dos sinais e sintomas apresentados pelos pacientes, tem a função de entender quais os mecanismos envolvidos, e quais são as alterações morfológicas e funcionais das células e tecidos doentes (KUMAR, ABBAS, e ASTER, 2013).

Durante o estudo da medicina e patologia, os 2 principais termos utilizados são etiologia e patogenia. Etiologia é a origem da doença, quais as causas e os fatores que geraram a doença. Patogenia é o desenvolvimento da doença, quais os fatores etiológicos iniciaram as alterações que originaram a doença, descreve como a doença se desenvolveu. A definição destes dois termos constitui a base para o tratamento, (KUMAR, ABBAS, e ASTER, 2013).

Para que seja possível diagnosticar a doença é necessário efetuar um estudo macro ou microscópico do tecido, células ou fluidos corporais (sangue e urina), isso é feito utilizando diversas técnicas morfológicas, moleculares, microbiológicas, estruturais e funcionais, para entender como os tecidos vão reagir em resposta a lesão, após todos os testes é possível orientar a terapia a ser utilizada (KUMAR, ABBAS, e ASTER, 2013).

2.2.1 Histopatologia tradicional

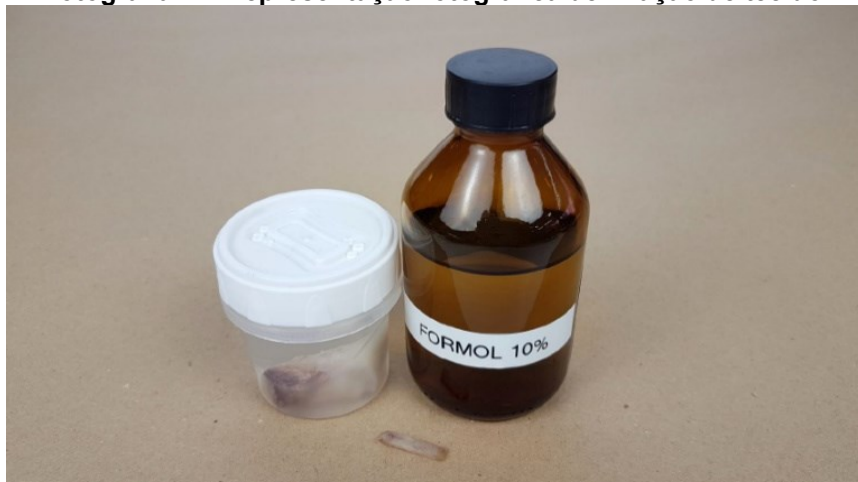
Tradicionalmente os exames histopatológicos são efetuados após a coleta das amostras de tecidos pela biopsia. Porém, para que possam ser analisados, os tecidos passam por um processo de preparação, para que seja possível analisá-los por microscopia (JUNQUEIRA e CARNEIRO, 2013).

2.2.2 Fixação

O processo de fixação se dá logo após a retirada do fragmento de tecido ou célula do corpo. Sua finalidade é evitar a deterioração do tecido por bactérias ou enzimas contidas no próprio tecido, evitar o endurecimento e preservar as composições e estruturas dos tecidos. A fixação pode ser efetuada por métodos físicos ou químicos. No método de fixação química o tecido é mergulhado em solução estabilizante de forma que as moléculas formem pontes com moléculas adjacentes, estas soluções

são chamadas de fixadores. Para que os fixadores se difundam por completo no tecido, o ideal é que sejam compostos por fragmentos pequenos, garantindo melhor penetração e preservação do tecido. A fotografia 1 representa o reagente usado na fixação da peça de tecido através de um banho de formol 10%, nestes casos o tempo varia dependendo do tamanho do tecido fixado (JUNQUEIRA e CARNEIRO, 2013; MORETTI et al., 2019).

Fotografia 1 – Representação fotográfica de fixação de tecido.



Fonte: MORETTI et al. (2019)

2.2.3 Desidratação

O processo de desidratação deve ser efetuado após o processo de fixação e tem a função de remover os líquidos e o formol utilizado no processo de fixação. Para que a desidratação seja possível, são necessário banhos em solução de concentração crescente de álcool, geralmente se inicia em álcool 70% até álcool 100% conforme representado na fotografia 2. Estes banhos são um banho de aproximadamente 1 hora para os álcoois de 70% até 90% e dois banhos no álcool a 100% (JUNQUEIRA e CARNEIRO, 2013; MORETTI et al., 2019).

Fotografia 2 – Representação fotográfica de desidratação da peça.



Fonte: MORETTI et al. (2019)

2.2.4 Diafanização

O processo diafanização é efetuado devido a necessidade de se substituir o álcool por uma substância intermediária que pode se misturar tanto em álcool quanto no meio escolhido para inclusão. O solvente utilizado geralmente é o xilol, comumente chamado de agente clarificador, por tornar o tecido transparente ou translúcido. A fotografia 3 representa o banho de xilol para diafanização, deixando a peça de tecido pronta para receber a inclusão (JUNQUEIRA e CARNEIRO, 2013; MORETTI et al., 2019).

Fotografia 3 – Representação fotográfica da Diafanização.



Fonte: MORETTI et al. (2019)

2.2.5 Inclusão ou Emblocamento

O processo de inclusão deve ser efetuado após todos os processos citados anteriormente e tem a função de enrijecer o tecido, as substâncias mais utilizadas neste processo são a parafina e algumas resinas plásticas, dependendo do tipo de microscopia (i.e., óptica e eletrônica). No caso da parafina, a temperatura usada para derretê-la (56-60 °C), permite que o solvente utilizado na diafanização evapore e os espaços existentes no tecido sejam preenchidos com parafina. A fotografia 4 representa o cassete de emblocamento do tecido em parafina, (JUNQUEIRA e CARNEIRO 2013; MORETTI et al. 2019).

Fotografia 4– Representação fotográfica do emblocamento do tecido em parafina.



Fonte: MORETTI et al. (2019)

2.2.6 Posicionamento e Microtomia

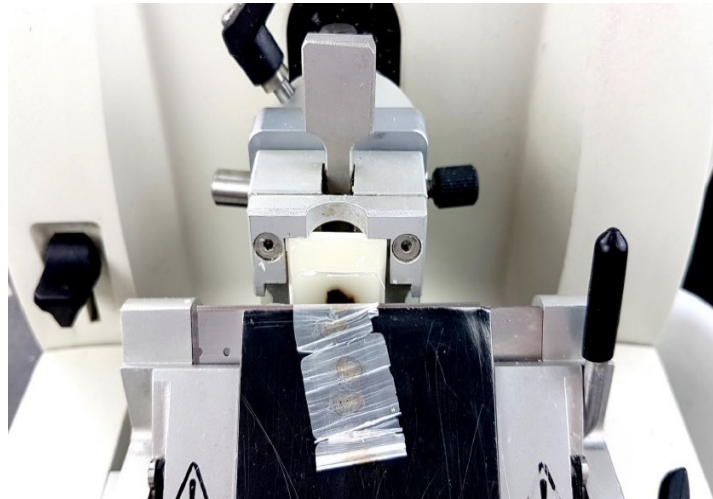
O processo de posicionamento e microtomia, é o processo de se posicionar o bloco de parafina com o tecido, colocá-lo em um aparelho chamado micrótomo, que faz o fatiamento ou seccionamento (i.e., corte histológico) com uma navalha descartável de aço ou vidro. O micrótomo fornece cortes de 1 a 10 micrômetros de espessura. A fotografia 5 representa o posicionamento de um bloco de parafina no micrótomo e a fotografia 6 representa o seccionamento do bloco para extração dos cortes com os tecidos (JUNQUEIRA e CARNEIRO, 2013; MORETTI et al., 2019).

Fotografia 5– Representação fotográfica do posicionamento de um bloco de parafina no micró-tomo.



Fonte: MORETTI et al. (2019)

Fotografia 6 – Representação fotográfica de um seccionamento de um bloco de parafina com tecido.

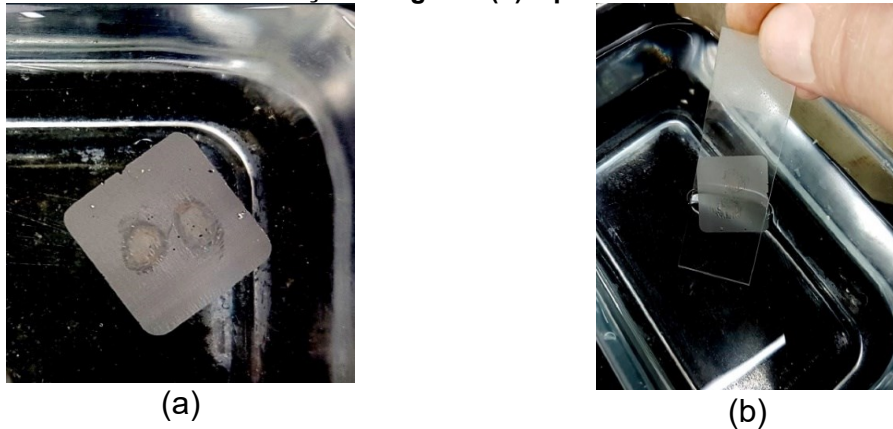


Fonte: MORETTI et al. (2019)

2.2.7 Separação e pesca dos cortes

O processo de separação e pesca ocorre após o corte histológico, onde o tecido é depositado dentro de um recipiente com água em temperatura ambiente, possibilitando o manuseio, separação dos tecidos e preparação para depositar na lâmina. O depósito na lâmina se dá ao encostar a lâmina no corte histológico efetuando a pesca, o tecido se adere ao vidro conforme o tecido é tracionado para fora da água. A fotografia 7(a) representa a separação dos tecidos dentro de um recipiente com água na temperatura ambiente, e a Fotografia 7(b) representa a pesca para levar ao banho de distensão da peça (MORETTI et al., 2019).

Fotografia 7 – Representação fotográfica da separação e pesca, fotografia (a) mostra a separação fotografia (b) a pesca.



Fonte: MORETTI et al. (2019)

2.2.8 Distensão e pesca dos cortes

O processo de distensão é semelhante ao processo de separação, porém, neste processo o tecido é depositado em banho a uma temperatura de aproximadamente 44°C, para que toda imperfeição ou dobra do tecido seja eliminada, para após ser depositado na lâmina após a pesca do corte. A fotografia 8 representa o processo de distensão e pesca do tecido (MORETTI et al., 2019).

Fotografia 8 – Representação fotográfica do processo de distensão e pesca do tecido.



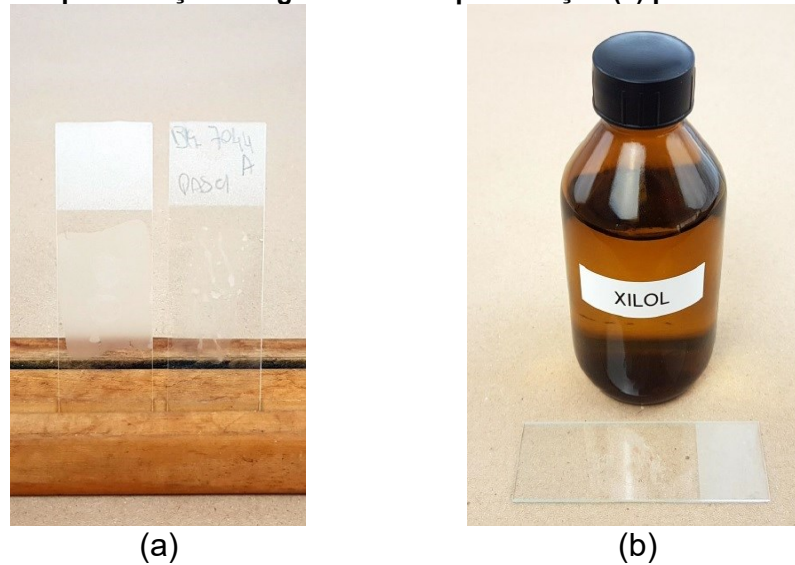
Fonte: MORETTI et al. (2019)

2.2.9 Desparafinação

O processo de desparafinação é efetuado para que se remova o excesso de

parafina contido no corte histológico e favorecer a aderência do corte na lâmina. Este processo é efetuado em estufa a 68°C, a lâmina é fixada em um suporte na posição vertical por aproximadamente 40 minutos. Após o processo de desparafinação pela estufa, o tecido passa por dois banhos de cinco minutos em xilol também com a intenção de remover o excesso de parafina. A fotografia 9(a) e (b), representam o processo de desparafinação por estufa e banho com xilol, (MORETTI et al., 2019; JUNQUEIRA e CARNEIRO, 2013).

Fotografia 9 – Representação fotográfica da desparafinação (a) por estufa e (b) por xilol.



Fonte: MORETTI et al. (2019)

2.2.10 Hidratação e Lavagem

O processo de hidratação tem como principal objetivo efetuar a limpeza do tecido, retirando o xilol, utilizado para desparafinação. Neste momento o tecido passa por 5 a 10 banhos de álcool na concentração decrescente de 100% até 70% de álcool. A lavagem é feita em água corrente para eliminar qualquer traço de álcool que o tecido possa conter, deixando o tecido pronto para receber a coloração. A fotografia 10 mostra a representação da hidratação e lavagem (JUNQUEIRA E CARNEIRO, 2013; MORETTI et al. 2019).

Fotografia 10 – Representação fotográfica da hidratação e lavagem, (a) banho de álcoois para hidratação e (b) lavagem.



Fonte: MORETTI et al. (2019)

2.2.11 Coloração

O processo de coloração dos cortes histológicos deve ocorrer devido a maior parte dos tecidos serem incolores para serem visualizadas em microscópios. Este método visa evidenciar os componentes contidos nos tecidos e nas células. Os corantes são substâncias químicas ácidas ou básicas e tendem a formar ligações eletrostáticas devido aos componentes ionizados dos tecidos. Basófilos e acidófilos são os componentes corados por afinidade com os corantes básicos e ácidos, respectivamente (JUNQUEIRA e CARNEIRO, 2013).

De todos os corantes utilizados para coloração de tecidos renais, a principal combinação utilizada é a hematoxilina com a eosina (HE). O corante com hematoxilina tem uma coloração azul ou violeta e cora as estruturas ácidas como núcleo celular, porções do citoplasma rica em ácido ribonucleico (RNA). O corante com eosina ao contrário da hematoxilina cora o colágeno e o citoplasma em cor rosa (JUNQUEIRA e CARNEIRO, 2013).

Inicialmente, a lâmina com o corte histológico é corada por um minuto em temperatura ambiente com o corante hematoxilina em uma determinada concentração. A lavagem ocorre após a coloração com a hematoxilina, para remover o excesso de corante e deve ser feita com água em abundância. A fotografia 11 representa a lâmina corada com hematoxilina e como a peça fica após o banho (MORETTI et al., 2019; JUNQUEIRA e CARNEIRO, 2013).

Fotografia 11 – Representação fotográfica do banho de hematoxilina para coloração das estruturas do tecido.



Fonte: MORETTI et al. (2019)

Após o processo de coloração com hematoxilina, o tecido passa pelo processo de diferenciação, esse processo é utilizado para melhorar e aumentar o contraste dos núcleos celulares. Ele é efetuado através de um mergulho rápido no diferenciador, uma solução de álcool ácido a 1%. Após o banho de álcool ácido, o processo de lavagem é efetuado novamente, para remover o excesso de diferenciador. Para aumentar a tonalidade azulada dos núcleos celulares, utilizasse água amoniacal. Ele é efetuado com um mergulho em solução de água amoniacal a 0,5%. Após o banho em água amoniacal, é efetuado um banho de água em abundância para eliminar o excesso da solução de água amoniacal do tecido. Após, é necessário o banho de álcool a 95% que tem como principal função remover qualquer resquício de água utilizada na lavagem das peças que ocorre após o banho de água amoniacal.

Após as lavagens com os solventes anteriores é feita a coloração com o corante eosina em uma determinada concentração, para a coloração, o tecido passa por um banho na eosina por aproximadamente um minuto, variável de acordo com a temperatura do ambiente onde o tecido está sendo corado. Três a cinco banhos de álcool a 100% são realizados com a função de remover o excesso de eosina do tecido. A fotografia 12 representa o banho de eosina para coloração das estruturas básicas do tecido (MORETTI et al., 2019).

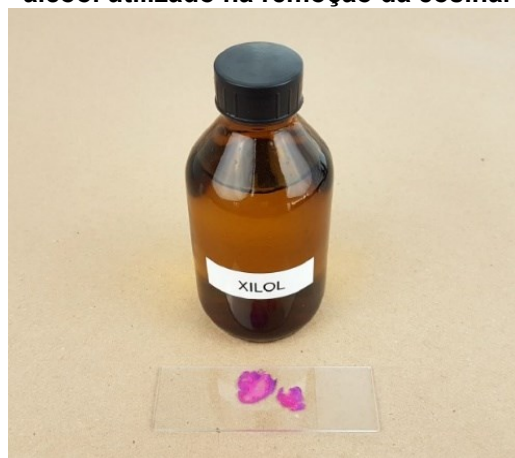
Fotografia 12 – Representação fotográfica do banho de eosina para coloração das estruturas básicas do tecido.



Fonte: MORETTI et al. (2019)

O processo de diafanização é realizado para remover o álcool 100% utilizado na remoção da eosina. A diafanização é feita com cinco a dez mergulhos em dois banhos em uma solução de xilol. A fotografia 13 representa o banho em xilol (MORETTI et al., 2019).

Fotografia 13 – Representação fotográfica dos banhos em solução de xilol, para remoção do álcool utilizado na remoção da eosina.

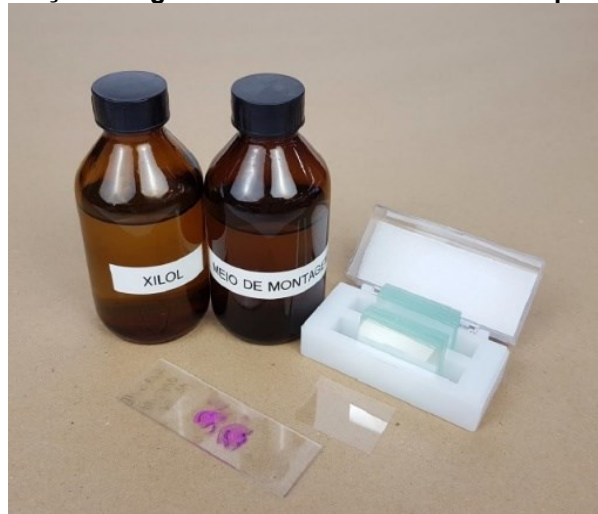


Fonte: MORETTI et al. (2019)

2.2.12 Montagem, limpeza, secagem e identificação

Após os processos de preparação das lâminas através da biopsia, desidratação, emblocamento, cortes histológicos, hidratação e lavagens, faz-se a secagem, limpeza e montagem da lâmina, seguida pela sua identificação. Inicialmente é necessário a separação do material que será utilizado como, meio de montagem, solução de xilol, lamínula (e.g., lâmina de vidro que cobre o tecido), lenço de papel e pinça. A fotografia 14 demonstra os materiais utilizados para montagem (MORETTI et al., 2019).

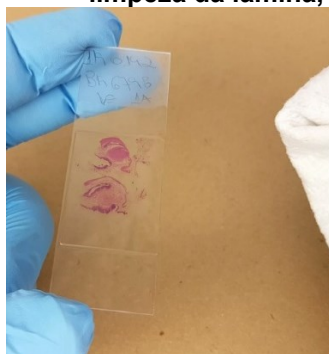
Fotografia 14 – Representação fotográfica dos materiais utilizados para montagem.



Fonte: MORETTI et al. (2019)

A montagem se dá através da inserção da lamínula sobre o corte histológico. Para isso é colocada uma gota do meio de montagem sobre a lâmina com o tecido e em seguida é posicionada a lamínula em cima da lâmina, de forma que a lamínula proteja o tecido, tornando essa lâmina permanente, após a montagem, é efetuado o processo de limpeza e etiquetagem para o correto armazenamento da lâmina. A limpeza é efetuada para retirar o excesso de meio de montagem que pode escorrer no processo de junção da lâmina com a lamínula. Para a etiquetagem que pode ser efetuada através de caneta, lápis ou uma etiqueta impressa, devemos esperar por aproximadamente 24 horas a secagem do meio de montagem. A fotografia 15 (a) e (b) representa o processo de limpeza e etiquetagem respectivamente, (MORETTI et al., 2019).

Fotografia 15 – Representação fotográfica da limpeza secagem e etiquetagem, (a) representa a limpeza da lâmina, (b) representa a secagem e etiquetagem.



(a)



(b)

Fonte: MORETTI et al. (2019)

2.3 Histopatologia digital

A histopatologia digital, passa pelos mesmos processos da patologia tradicional sendo incluída mais uma etapa que é a digitalização da lâmina (JUNQUEIRA e CARNEIRO, 2013; MORETTI et al., 2019). Com o avanço tecnológico, as melhorias em hardware e software. A microscopia convencional vem dando lugar a histopatologia digital, um novo conceito, uma quebra de paradigma, que modifica a forma de análise médica, que anteriormente ficava restrito a visualização parcial de uma pequena área do tecido devido à área de foco dos equipamentos de microscopia (DOUGHERTY, 2009).

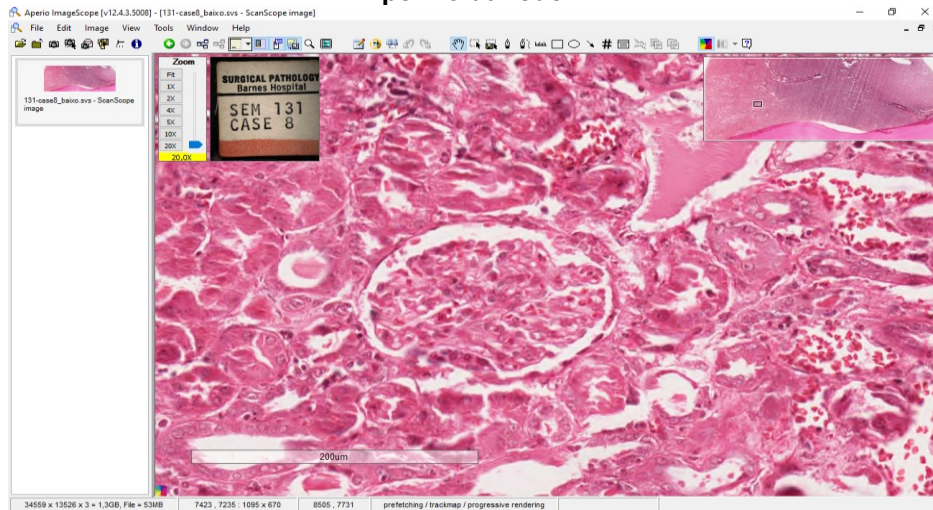
Na histopatologia digital, as lâminas de tecido são digitalizadas para o tamanho total da lâmina em alta resolução. Estas imagens podem ser armazenadas em um servidor de arquivos ou banco de dados (BD), permitindo assim que as imagens sejam compartilhadas de maneira simples e fácil entre os médicos. Permite também o compartilhamento com a comunidade científica, possibilitando estudos e melhorias nas ferramentas já existentes, (DOUGHERTY, 2009).

Técnicas de computação gráfica podem aumentar a sensibilidade das ferramentas, permitindo desta forma que os laudos patológicos tenham informações com maior exatidão, isso é possível na histopatologia digital porque as imagens são formadas por uma matriz de pixel o que torna possível a identificação de formas, contagem de objetos entre outros (OBERHOLZER et al., 1996).

Imagens de lâminas inteiras estão sendo amplamente utilizadas na patologia digital e a sua utilização tem crescido cada vez mais, pois tem mostrado melhor que outros processos de diagnósticos que não podem contar com técnicas de processamento de imagens, (BUENO et al., 2016).

As imagens visualizadas ao microscópio óptico em ampliações variadas (e.g., 10X, 20X, 40X, 100X) pelos médicos podem ser transferidas para os monitores de alta definição, permitindo a visualização de uma maior área do tecido. A utilização de softwares que possibilitam a navegabilidade através das imagens, zoom, entre outras funcionalidades, permitindo assim maior exatidão no diagnóstico médico. A figura 4 mostra um exemplo da tela do software Aperio ImageScope utilizado na pesquisa para fazer o download das imagens utilizadas no processo de treinamento, validação e análise de desempenho da rede (BUENO et al., 2016), também podemos utilizá-lo como um exemplo de software de manipulação de lâminas digitalizadas inteiras.

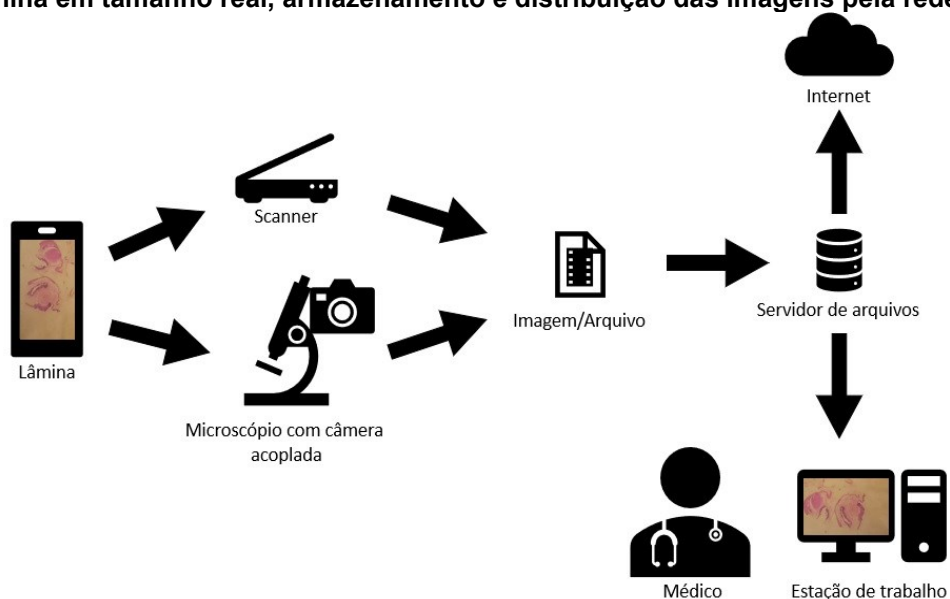
Figura 4 – Tela de exemplo do software Aperio ImageScope utilizado na pesquisa para fazer o download das imagens utilizadas no processo de treinamento, validação e análise de desempenho da rede.



Fonte: Autoria própria (2020)

Este software pode ser baixado gratuitamente através da página da empresa Leica, este software não tem somente a funcionalidade de download, com ele se pode efetuar diversas análises como ajustar a ampliação, zoom, comparação, efetuar anotações em áreas de interesse, entre outras funcionalidades (Leica 2020). A figura 5 demonstra o diagrama em bloco da lâmina finalizada, o escaneamento ou fotografia sequencial da imagem, conversão da imagem para formato digital, o armazenamento em servidor de arquivo e a distribuição que pode ser pela internet ou intranet.

Figura 5 – Representação em blocos da lâmina montada e do processo de digitalização da lâmina em tamanho real, armazenamento e distribuição das imagens pela rede.



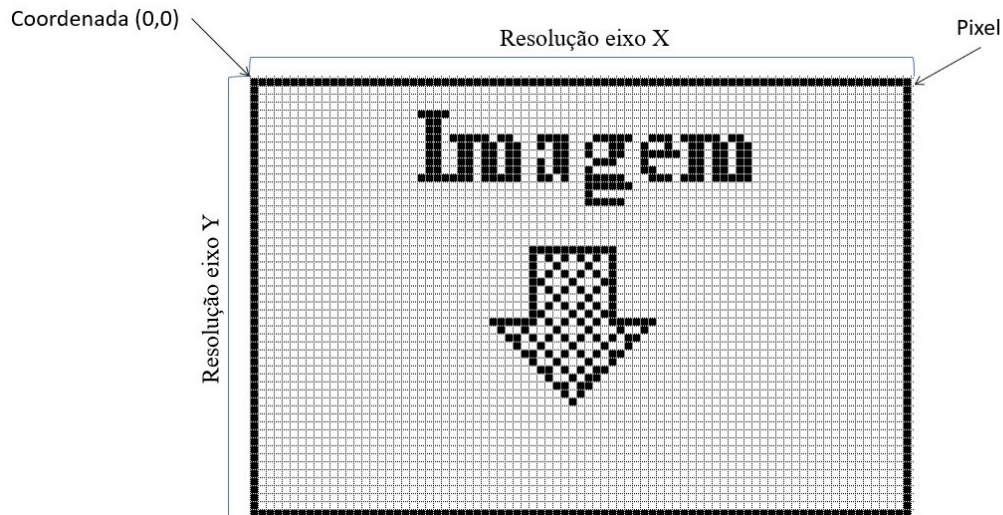
Fonte: Autoria própria (2020)

2.4 Imagem

Imagem é uma representação discreta dos dados obtidos através de equipamentos como máquinas fotográficas analógicas, câmeras digitais, Raio X, tomógrafos. Ela pode ser representada em uma forma física (e.g., papéis fotográficos, filmes radiográficos) ou em formato digital, salvas em computadores ou memórias de câmeras digitais. Há diversos formatos de imagens tais como, GIF (Formato para intercambio de imagem), JPG (Grupo Misto de Especialistas Fotográfico), BMP (Imagem em mapa de bit), PNG (Imagens portáteis para redes), TIF/TIFF (Formato de imagem com rótulo/ Formato de arquivo de imagem com rótulo) (SOLOMON e BRECKON, 2013).

As imagens digitais em tons de cinza têm uma grande quantidade de informação definida como um sinal espacial, representado pela função $f(x, y)$, onde x e y são as coordenadas e f em qualquer par de coordenadas (x, y) é o nível de cinza. Desta forma podemos dizer que a imagem digital é formada por um número finito de elementos e que a sua localização nas coordenadas (x, y) pode ser chamado de pixel. A figura 6 mostra como uma imagem em tons de cinza é formada (GONZALEZ e WOODS, 2009).

Figura 6 – Representação de uma imagem com resolução de 83 x 55 pixels, o ponto de início das coordenadas (0,0) e a representação de um pixel.



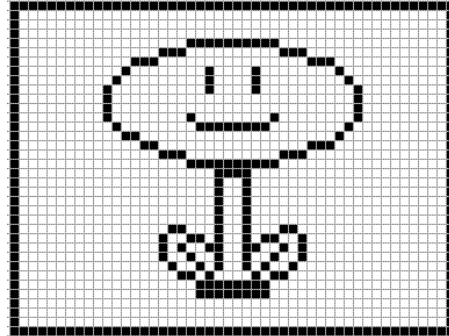
Fonte: Autoria própria (2020)

2.4.1 Cor em imagem digital

A imagem pode conter um ou mais canais, estes canais são responsáveis

pela definição e intensidade das cores em um pixel $f(x, y)$ (SOLOMON e BRECKON, 2013). Nas imagens mais simples, cada pixel é formado por um único valor ou nível de sinal, como exemplo podemos utilizar uma imagem em preto e branco, formando a imagem com zeros e uns. A figura 7 mostra uma imagem em preto e branco, (SOLOMON e BRECKON, 2013).

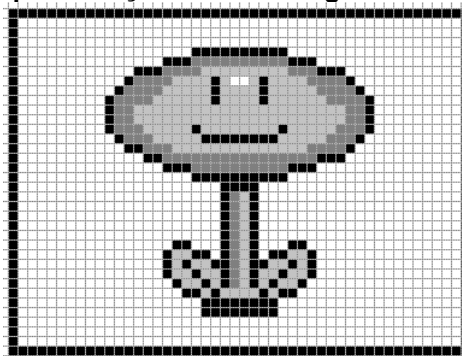
Figura 7 – Representação de uma imagem em preto (0) e branco (255).



Fonte: Autoria própria (2020)

As imagens também podem ser representadas por um mapa de cor, onde cada pixel pode alocar um tom específico de cor para representar a imagem, o mapa de cor mais simples e comum é a escala de cinza que pode alocar os valores de preto (0) até branco (255). A figura 8 representa uma imagem em mapa de cor em escala de cinza (SOLOMON e BRECKON, 2013).

Figura 8 – Representação de uma imagem em escala de cinza.

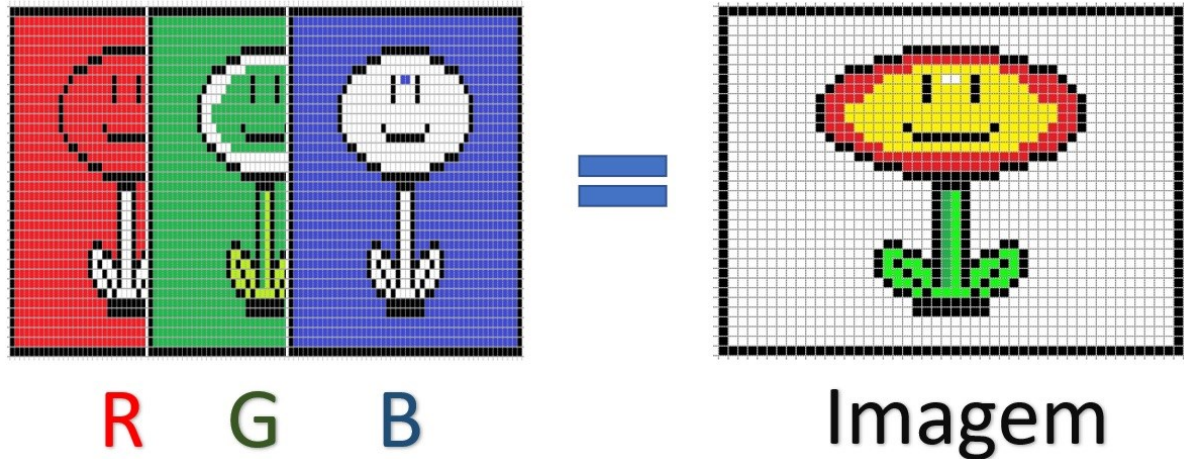


Fonte: Autoria própria (2020)

Além das imagens em escala de cinza, onde cada pixel recebe um valor numérico, temos imagens com cores reais podendo conter todo o espectro de cores, e o pixel pode ser representado como um vetor com três componentes de cores (R, G, B), Red ou vermelho, Green ou verde e Blue ou azul. Podemos chamar também de camadas RGB. A figura 9 representa como é formada a imagem através da junção

das camadas de cores, (SOLOMON e BRECKON, 2013).

Figura 9 – Representação de como é formada uma imagem digital com camadas RGB.



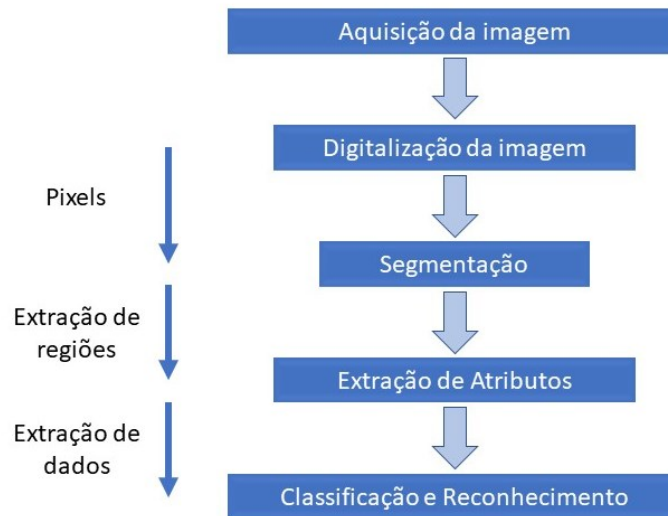
Fonte: Autoria própria (2020)

Outra representação de cor amplamente utilizada é o mapa de cor (H, S, V) que corresponde às componentes Hue (i.e., matiz, tonalidade), Saturation (i.e., saturação) e Value (i.e., valor, intensidade), onde a intensidade é separada da informação de tonalidade e saturação (SOLOMON e BRECKON, 2013).

2.4.2 Processamento digital de imagem

O processamento digital de imagem (PDI) é a manipulação de uma imagem digital com a aplicação de algoritmos computacionais e tem como objetivo melhorar o aspecto visual de forma a realçar, evidenciar ou destacar objetos ou estruturas nas imagens, facilitando assim identificação e extração de informações. Dependendo do tipo de processamento que será efetuado, é necessário a utilização de hardwares específicos, que auxiliam no processamento das imagens, como placas de vídeo. A área de processamento de imagens aumentou significativamente atualmente, e as suas aplicações estão em diversas atividades do dia a dia, como no geoprocessamento agrícola de precisão para análise de lavouras, na segurança pública para monitoramento através de câmeras, no controle de qualidade de empresas, nos sistemas de medição, na medicina diagnóstica, entre outros ramos de atividade. A figura 10 mostra o diagrama em bloco de uma estrutura básica de processamento de imagem (NETO e FILHO, 1999).

Figura 10 – Representação em blocos dos passos para processamento digital de imagens.



Fonte: Autoria própria (2020)

2.4.2.1 Aquisição da imagem

Os dispositivos de aquisição de imagens variam desde uma simples câmera digital ou até mesmo um satélite, equipamentos de ressonância magnética entre outros (NETO e FILHO, 1999).

2.4.2.2 Digitalização da imagem

A etapa de digitalização da imagem tem a função de converter um sinal analógico em código digital, através de equipamento (i.e., Scanners, máquinas fotográficas) e softwares (e.g., sistema de gerenciamento eletrônico de documentos) (NETO e FILHO, 1999).

2.4.2.3 Segmentação

A Segmentação tem o objetivo de extrair ou enfatizar a área que representa o objeto de interesse. Isso pode ser efetuado através da segmentação da imagem original em uma ou mais imagens (NETO e FILHO, 1999).

2.4.2.4 Extração de atributos

A extração de atributos tem a finalidade de extrair as características das imagens resultantes do processo de segmentação e pós-processamento, diferenciar e evidenciar as similaridades entre os objetos. A aparência visual geralmente define as

características do objeto, como aspecto, cor, tamanho e perímetro, forma (NETO e FILHO, 1999).

2.4.2.5 Classificação e reconhecimento

Classificação e reconhecimento tem o objetivo de reconhecer os padrões e classificar objetos através de algoritmos computacionais podemos dizer que é a capacidade de um computador de tentar imitar a capacidade humana de reconhecer os padrões. Chamada de visão computacional, técnica que permite que computadores através de algoritmos computacionais possam enxergar, reconhecendo placas, peças, caracteres, faces humanas e outros objetos (NETO e FILHO, 1999).

2.4.3 Processo humano de identificação

Estima-se que um ser humano utiliza entre 90% e 95% da sua visão para adquirir as informações. Podemos dizer que nós seres humanos somos seres mais visuais, a nossa visão recebe a maior parte das informações externas, embora tenhamos outros sentidos como tato, olfato, audição ou paladar. Embora este seja nosso principal sentido, tenhamos milhões de anos de evolução este nosso sentido não lida muito bem com a compreensão das diferenças dos tipos de informação a serem extraídas das imagens. Para que a visão humana funcione com eficiência, dependemos da nossa evolução e da experiência, porém muitas vezes algumas informações não são observadas ou são suprimidas, ignoradas. Geralmente somos capazes de detectar somente coisas esperadas ou coisas que estamos concentrados para encontrar (RUSS e NEAL, 2016).

2.4.3.1 Reconhecimento

O objetivo da visão humana em sua maior parte é o reconhecimento de objetos, geralmente quando observamos uma imagem, o que nos chama a atenção é algo familiar que já conhecemos, isso ocorre porque temos em nossa mente objetos com características vinculadas a nomes (rótulos), que rapidamente nossa consciência faz a atribuição, atrás destes rótulos temos o modelo mental do objeto que pode ser em forma de lembranças, imagens, palavras, eventos, experiências vividas, entre outros. Porém a dificuldade não está em reconhecer objetos familiares e sim no reconhecimento de objetos desconhecidos, coisas que não temos um nome ou um rótulo

definido, para estas coisas temos uma maior dificuldade em reconhecer. Cada pessoa tem seu exemplo formado, a sua representação armazenada, suas memórias individuais. Por este motivo sempre que é necessário solicitar algum objeto, devemos especificar de maneira clara o que é solicitado, caso contrário abrimos caminhos para falhas de comunicação (RUSS e NEAL, 2016).

2.4.3.2 Processamento de imagem de um olho humano

O processamento de imagem da visão humana depende de outras estruturas além das que os olhos contêm. Parte deste processamento ocorre imediatamente na retina e parte ocorre no cérebro, dentro do que é chamado de córtex visual, localizado na parte posterior do cérebro.

2.5 Inteligência artificial

Inteligência artificial é uma área da ciência da computação que tem como finalidade simular a capacidade humana de resolver tarefas e problemas, através de algoritmos computacionais, que vão desde simples programas especialistas até algoritmo de aprendizado de máquina ou aprendizado profundo, algoritmos que aprendem de forma independente. IA teve seu início na década de 40. Uma das primeiras publicações para a comunidade científica foi o artigo *Computing Machinery and Intelligence* no ano de 1956 (TURING, 1950), do matemático britânico Alan Mathison Turing, logo após a Segunda Guerra Mundial, conhecido como pai da computação por ter tido influência no desenvolvimento da ciência da computação, na formalização do conceito de algoritmos, foi um dos pioneiros na IA e na ciência da computação (TAULLI, 2020).

O campo de aplicações de IA está em franca expansão devido o avanço tecnológico dos sistemas computacionais (KREUTZER e SIRRENBURG, 2020). IA está em busca não só de compreensão, mas sim de construir entidades inteligentes, que possam calcular como agir com segurança e eficácia em uma gama de situações novas (RUSSELL e NORVIG, 2020). Na literatura, temos pequenas diferenças em como Inteligência Artificial é definida, mas em resumo ela pode ser explicada de 4 maneiras (RUSSELL e NORVIG, 2020):

- Sistemas que pensam como humanos;
- Sistemas que pensam racionalmente;

- Sistemas que agem como humanos;
- Sistemas que agem racionalmente.

2.5.1 Aprendizado de máquina

Aprendizado de máquina é uma subárea de IA, é a ciência da computação que dá aos computadores a habilidade de aprender com os dados. Os algoritmos utilizados no aprendizado de máquina, são algoritmos especiais chamados de algoritmos auto adaptativos que permitem o aprendizado de forma independente sem intervenção de um programador no processo de aprendizagem (KREUTZER e SIRRENBORG, 2020). Para que tenhamos um bom processo de treinamento e aprendizado, para o problema a ser tratado, necessitamos de uma grande quantidade de dados com qualidade. Estes dados podem ser revisados e retreinados a fim de melhorar a base para as tomadas de decisão (KREUTZER e SIRRENBORG, 2020).

Podemos diferenciar o aprendizado de máquina em 3 tipos, aprendizado supervisionado, aprendizado não supervisionado e aprendizagem por reforço.

- **Aprendizado supervisionado** – Neste processo se aprende uma função com base na observação dos pares, entrada e saída. Característica básica é que neste modo os dados utilizados para treinamento contêm as respostas desejadas, exemplo disso é, o treinamento da rede é efetuado com os dados anotados com as respostas às classes a serem previstas (KREUTZER e SIRRENBORG, 2020 e RUSSELL e NORVIG, 2020).
- **Aprendizado não supervisionado** – Neste processo o aprendizado ocorre sem nenhuma informação de anotação explícita, o usuário não tem ciência dos padrões, o algoritmo deve reconhecer independentemente uma estrutura. (KREUTZER e SIRRENBORG, 2020 e RUSSELL e NORVIG, 2020).
- **Aprendizado por reforço** – Neste processo o aprendizado ocorre através de recompensas e punições, o sistema experimenta iterativamente de forma independente por meio de um processo de tentativa e erro, descartando ou absorvendo, através de recompensas para bons resultados ou punições para resultados incorretos, esta abordagem geralmente é utilizada quando não se tem o resultado ideal, ou quando é necessário extrair

informação com a interação com o meio ambiente (KREUTZER e SIRRENBORG, 2020 e RUSSELL e NORVIG, 2020).

2.5.2 Redes neurais artificiais

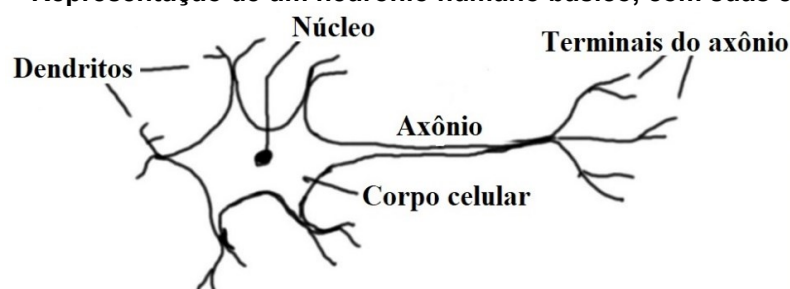
O Cérebro humano tem a capacidade de processar uma grande quantidade de dados em um tempo muito baixo. Os neurônios são as principais estruturas do cérebro, é no neurônio que ocorre o processamento e transmissão das informações. O que se sabe é que o cérebro humano desenvolve suas regras através de conhecimentos adquiridos em situações ocorridas anteriormente. Com base nisso, diversos pesquisadores vêm buscando formas de simular o funcionamento do cérebro, através do processo de aprendizagem por experiência, com o foco de criar sistemas inteligentes e capazes de reconhecer padrões, classificar objetos, processar imagens entre outras. Como resultado destas pesquisas, tivemos a criação de um modelo de neurônio artificial, que interconectados criam a chamada rede neural.

As redes neurais artificiais, surgiram junto com IA também na década de 40, mais especificamente em 1943 pelo matemático e o neurofisiologista Walter Pitts e Warren McCulloch (PITTS & MCCULLOCH, 1943) através de uma analogia efetuada entre as células nervosas e circuitos eletrônicos.

2.5.2.1 Modelo de um neurônio humano

A rede neural (NN), ou como chamada em algumas literaturas rede neural artificial (ANN) tem este termo originado da neurociência, emula o funcionamento de um neurônio artificial onde eles se conectam entre si, formando uma rede para processar os sinais químicos e elétricos, o núcleo processa as informações recebidas pelos dendritos, e o axônio é usado como um cabo para enviar informações. A figura 11 demonstra a estrutura básica de um neurônio humano.

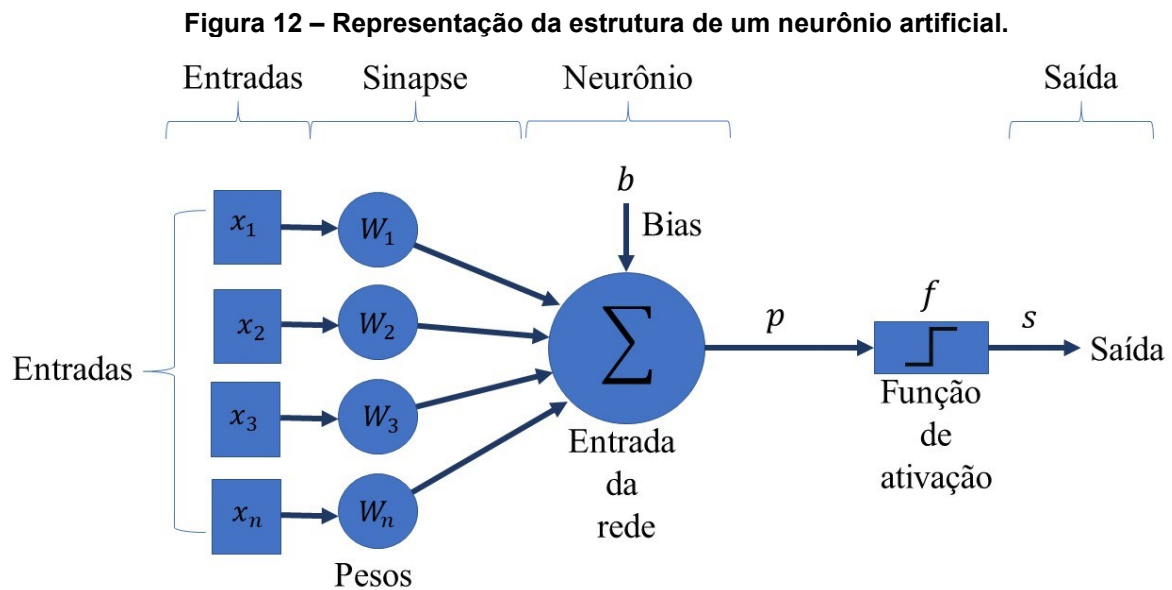
Figura 11 – Representação de um neurônio humano básico, com suas estruturas.



Fonte: SILAPARASETTY (2020, pg. 50)

2.5.2.2 Modelo de um neurônio artificial

O neurônio artificial em sua forma mais básica, é um modelo matemático criado com base em um neurônio biológico. A figura 12 mostra a estrutura de um neurônio artificial com seus componentes.



Fonte: Autoria própria (2020)

Como citado anteriormente, os neurônios biológicos são células nervosas compostas de três principais segmentos, os dendritos, o corpo celular e o axônio. Dendritos recebem os sinais através das junções chamadas de sinapses neurais, o corpo celular é responsável pelo processamento dos sinais recebidos pelos dendritos e um potencial de ativação indicará se a propagação de um impulso elétrico vai ou não ocorrer em direção ao axônio.

O neurônio artificial ou neurônio matemático funciona da seguinte forma. Sinais de entrada (x_i), são sinais externos que alimentam o modelo com um conjunto de informações do processo que estamos mapeando, representado pelo vetor $x = [x_1, x_2, x_3, x_n]$, são somadas e ponderadas pelos pesos sinápticos (w_i), representados pelo vetor $w = [w_1, w_2, w_3, w_n]$. O resultado da somatória (i.e., Σ) das entradas ponderadas é somado ao limiar de ativação (b), que é enviado como argumento para a função de ativação (f), no qual teremos a saída (s) desejada. A equação 1 demonstra a função de ativação. A Equação 2 representa o potencial de ativação.

$$s = f(p) \quad (1)$$

$$p = b + \sum_{i=1}^n w_i \cdot x_i \quad (2)$$

2.5.2.3 Funções de ativação

Uma função de ativação é uma transformação com comportamento linear e não linear, utilizada para decidir se o neurônio deve ou não ser ativado, ou seja, determina a relevância da informação que está sendo recebida, se ela é ou não relevante para a saída (SILAPARASETTY, 2020).

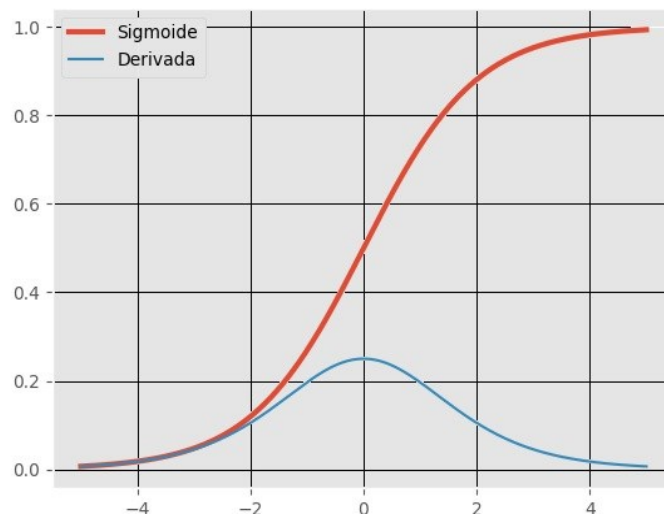
Abaixo, descrevo os diferentes tipos de funções de ativação, suas equações matemáticas e seus gráficos:

- Função de ativação Sigmoide ou logística: Função de ativação amplamente utilizada, a Função de ativação Sigmoide, assume um intervalos de valores entre 0 e 1. A equação 3 demonstra a função sigmoide, a equação 4 sua derivada e a Figura 28 demonstra o gráfico obtido ao aplicar as equações 3 e 4 (HAYKIN, 2003).

$$f(x) = \sigma(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

$$f'(x) = f(x)(1 - f(x)) \quad (4)$$

Figura 13 – Representação do gráfico gerado com a aplicação das Equações 3 e 4.



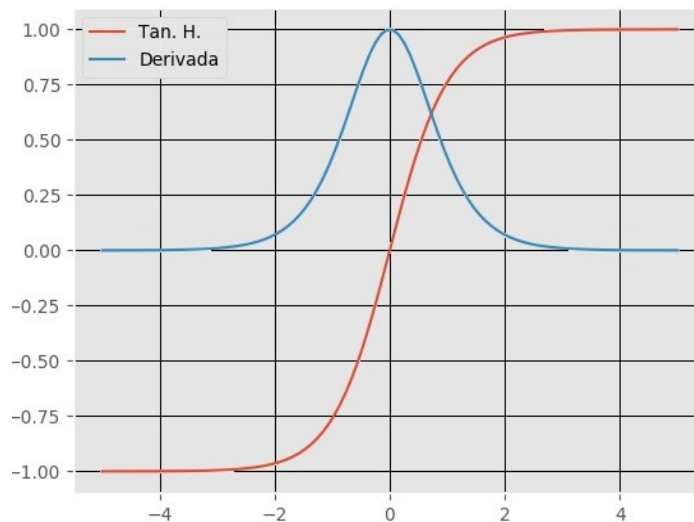
Fonte: Adaptado de FACURE, 2017 e SILAPARASETTY (2020)

- Função de ativação TanH (Tangente Hiperbólica): A função de ativação TanH é uma função semelhante à função Sigmoide, tem um formato em “S”, podendo variar entre o valor de 1 até -1. A equação 5 representa a função TanH, a equação 6 representa sua derivada, e a figura 14 representa o gráfico criado pelas equações 5 e 6 (SILAPARASETTY, 2020).

$$f(x) = \tanh(x) = \sigma(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})} \quad (5)$$

$$f'(x) = 1 - f(x)^2 \quad (6)$$

Figura 14 – Representação do gráfico gerado com a aplicação das equações 5 e 6.



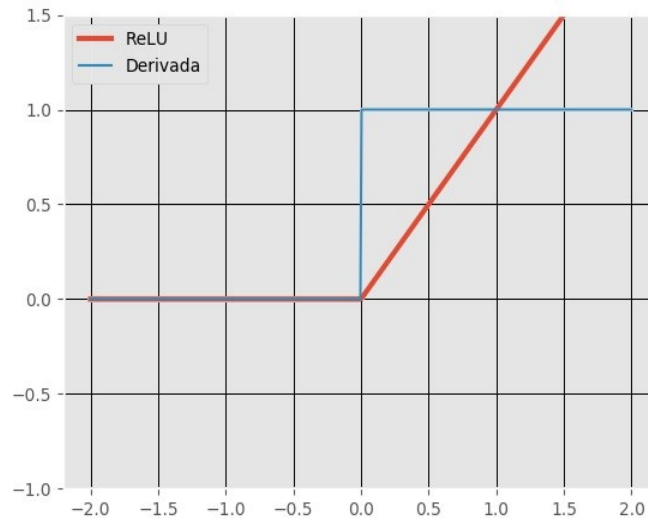
Fonte: Adaptado de FACURE, 2017 e SILAPARASETTY (2020)

- Função de ativação ReLU: ReLU ou Rectified Linear Unit (Unidade Linear Retificada) é uma função de ativação muito popular no aprendizado profundo. Como seu resultado para valores negativos, tende a desativar alguns neurônios durante um passo de *forward*, aumentando a velocidade do treinamento, podemos dizer que a função ReLU é mais eficiente computacionalmente na convergência da rede no momento do treinamento, que as funções Sigmoide e TanH. A equação 7 demonstra a função, a equação 8 sua derivada, e a figura 15 demonstra os gráficos gerados pelas equações 7 e 8 (SILAPARASETTY, 2020).

$$f(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ x & \text{se } x \geq 0 \end{cases} = \max\{0, x\} = x1_{x>0} \quad (7)$$

$$f'(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ 1 & \text{se } x > 0 \end{cases} \quad (8)$$

Figura 15 – Representação do gráfico gerado com a aplicação das equações 7 e 8.



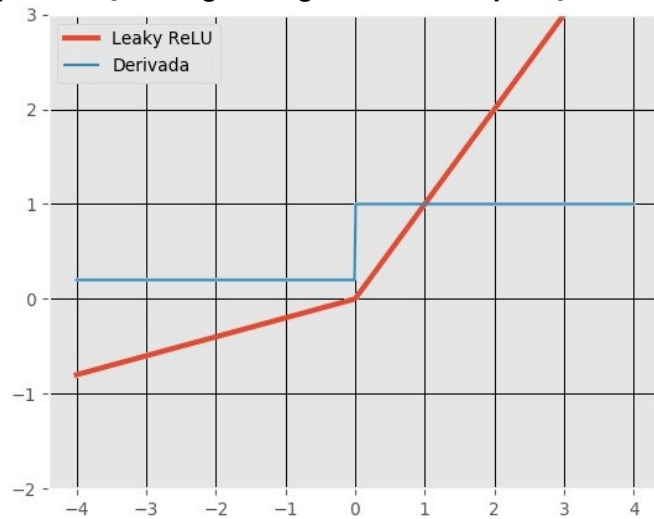
Fonte: Adaptado de FACURE, 2017 e SILAPARASETTY (2020)

- **Função de ativação Leaky ReLU:** A função de ativação Leaky ReLU, é uma atualização da versão ReLU que resolve o problema da função ReLU, ou seja, quando $x < 0$, a saída será sempre 0. Já na Leaky ReLU ao invés de definir 0 para valores negativos, (Maas et al, 2013) propôs uma pequena inclinação para a função na parte negativa. Com esta pequena modificação o lado esquerdo do gráfico passa a ser diferente de zero, desta forma não existirá mais neurônios desativados. A equação 9 demonstra a função, a equação 10 demonstra a derivada, e a figura 16 demonstra o gráfico gerado pelas equações 9 e 10 (SILAPARASETTY, 2020).

$$f(x) = \begin{cases} 0.01x & \text{se } x < 0 \\ x & \text{se } x \geq 0 \end{cases} \quad (9)$$

$$f'(x) = \begin{cases} 0.01 & \text{se } x < 0 \\ 1 & \text{se } x \geq 0 \end{cases} \quad (10)$$

Figura 16 – Representação do gráfico gerado com a aplicação das equações 9 e 10.

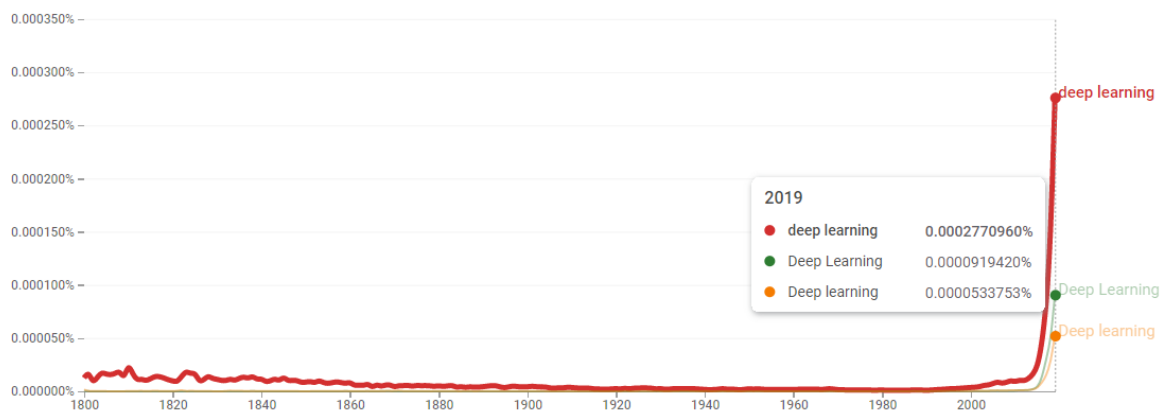


Fonte: Adaptado de FACURE, 2017 e SILAPARASETTY (2020)

2.5.3 Aprendizado profundo

O aprendizado profundo do inglês *Deep Learning*, é uma técnica de aprendizado de máquina (*Machine Learning*). A primeira publicação sobre o tema à comunidade científica foi em 1989 com o artigo intitulado *Backpropagation applied to handwritten zip code recognition* (LECUN et al., 1989), é possível ver que muitos conceitos surgiram antes de 2012, porém, como é possível ver na figura 17, gerada do site *Google Books Ngram Viewer*, a pesquisa da palavra *Deep learning* começou a ganhar a devida atenção na internet a partir 2012, com o artigo intitulado *Deep Learning Made Easier by Linear Transformations in Perceptrons* (RAIKO et al., 2012).

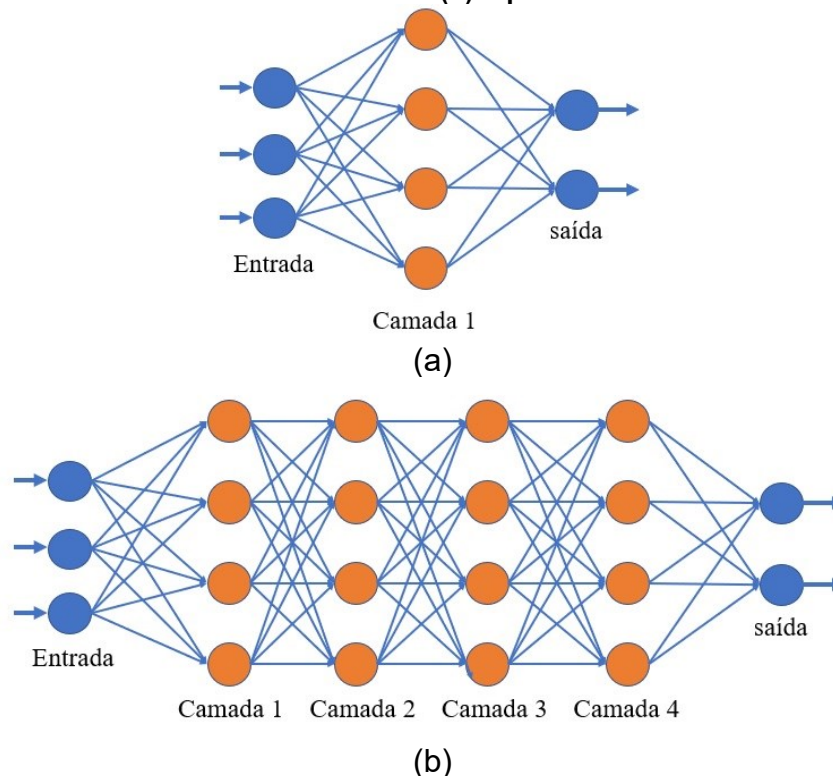
Figura 17 – Representação gráfica do aumento de popularidade da pesquisa do termo Deep Learning no Google Books Ngram Viewer, nos Estados Unidos.



Fonte: Google Books Ngram Viewer (2020)

Aprendizado profundo é uma técnica de aprendizado de máquina que possui uma configuração especial de rede neurais que tem a característica de poder processar uma gama mais ampla de dados, com menos pré-processamento de dados, geralmente apresentando resultados mais exatos do que as abordagens tradicionais de aprendizado de máquina. A palavra “profundo” representa a aprendizagem de ponta a ponta (i.e., *End to End (E2E)*), aprendizagem complexa ou não, representado por um modelo de aprendizado por característica (i.e., *Feature Learning*), um conjunto de técnicas para descobrir, de forma automática, as características necessárias para detecção ou classificação através de dados brutos. As redes neurais estão organizadas em muitas camadas, computacionalmente podemos falar que, da entrada até a saída, temos muitas etapas de processamento. A figura 18 mostra uma rede neural simples, de uma camada, e uma rede neural profunda com quatro camadas (KREUTZER e SIRRENBORG, 2020 e RUSSELL e NORVIG, 2020).

Figura 18 – Representação gráfica de rede neural simples e de múltiplas camadas, (a) representa uma rede neural de uma camada e (b) representa uma rede neural de 4 camadas.



Fonte: Autoria própria (2020)

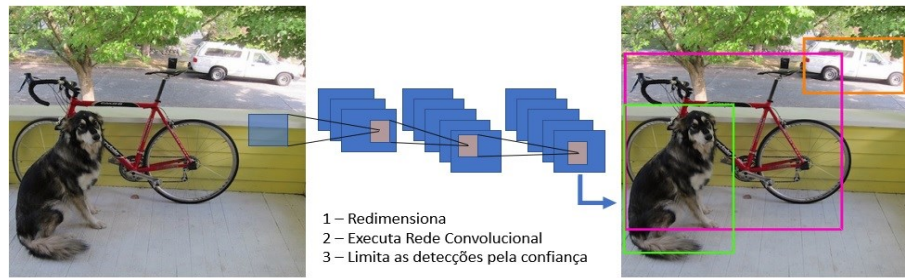
2.5.4 YOLOv3

You Only Look Once – YOLO (i.e., “Você só olha uma vez”) é uma arquitetura de rede neural de detecção unificada de objetos em imagens em tempo real, de vídeos, à 45 quadros por segundo, trata-se de uma ferramenta de código aberto, que permite o desenvolvimento colaborativo e de não ser necessária a compra de uma licença para utilização. As suas principais vantagens são que a arquitetura analisa a imagem inteira, desta forma ela considera todo o contexto da imagem sendo uma rede que pode ser otimizada de ponta a ponta devido sua estrutura unificada (REDMON et al. 2016). YOLO foi apresentado como uma nova abordagem para detecção de objetos. Trabalhos anteriores detectavam e classificavam as imagens em processos distintos. Devido a sua estrutura unificada, o YOLO é extremamente rápido, trata classificação e detecção de objetos como um problema de regressão, YOLO raciocina de forma global em toda a imagem para fazer as previsões (REDMON et al. 2016).

Uma abordagem recente como a R-CNN (Region-based convolutional neural network) é considerada um método superior, porém, o R-CNN utilizam primeiramente métodos de propostas de região, para depois gerar as caixas delimitadoras em potencial, nesta caixa delimitadora em potencial, é executado um classificador, em seguida, para eliminar possíveis erros de detecção duplicada e pontuar novamente as caixas é executado um pós-processamento, isso acaba tornando estas redes complexas e potencialmente lentas (REDMON et al. 2016).

O YOLO aprende de forma generalizável a representação dos objetos isso faz com que, quando treinado e testado, supere alguns dos principais métodos como DMP (Data and Model Parallel) e R-CNN. A figura 19 demonstra a complexidade do YOLO, após treinada a arquitetura “olha” apenas uma vez para a imagem para prever quais objetos estão presentes (REDMON et al. 2016).

Figura 19 – Representação do sistema de funcionamento YOLO, (1) redimensiona a imagem na entrada da rede, (2) executa uma única rede convolucional na imagem, (3) delimita as detecções resultantes pela confiança.



Fonte: Autoria própria (2020)

A unificação dos componentes de detecção de objetos em uma única arquitetura com rede neural, faz com que a rede raciocine globalmente sobre a imagem para a previsão das caixas delimitadoras, o que possibilita um treinamento mais rápido. O sistema divide a imagem de entrada em uma grade $N \times N$ onde cada célula da grade prevê 3 caixas delimitadoras em escalas diferentes B , para compensação da caixa delimitadora CC , previsão de objetividade PO e previsão de classes C (REDMON e FARHADI, 2018), de modo que o número de neurônios na camada de saída é dado pela equação 11.

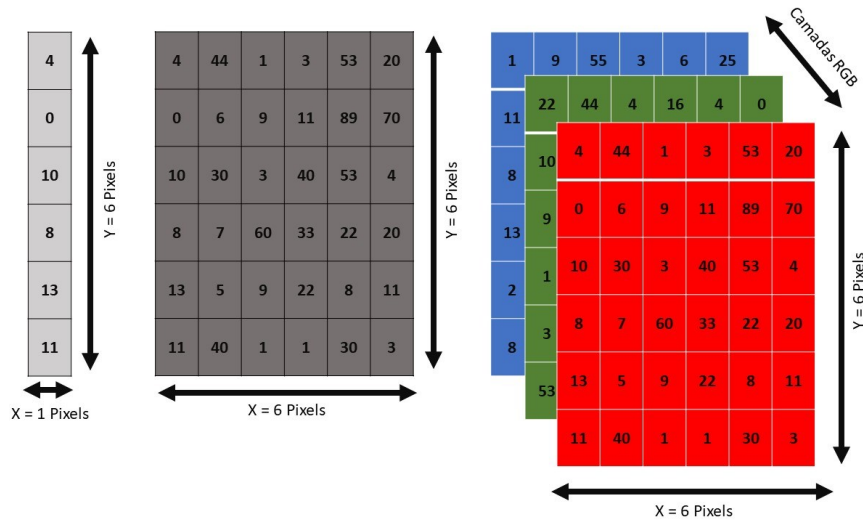
$$N_N = n \times n \times (b \times (cc + po + c)) \quad (11)$$

Para as informações contidas inicialmente nesta pesquisa, vamos adotar os seguintes valores, B igual a 3, CC igual a 4, PO igual a 1, C igual a 1, desta forma, a fórmula para o cálculo do número de neurônios na camada saída, ficaria da seguinte forma:

$$N_N = n \times n \times (3 \times (4 + 1 + 1))$$

Tensor pode ser definido como uma generalização de um vetor, onde o vetor seria um tensor unidirecional ou de primeira ordem ou uma matriz é um tensor bidirecional ou de segunda ordem, e uma matriz com três camadas é um tensor tridimensional (e.g., uma imagem colorida RGB) ou de terceira ordem, pode ser facilmente compreendido como um array multidimensional. A figura 20 demonstra e exemplifica o tensor unidirecional, tensor bidimensional e um tensor tridimensional.

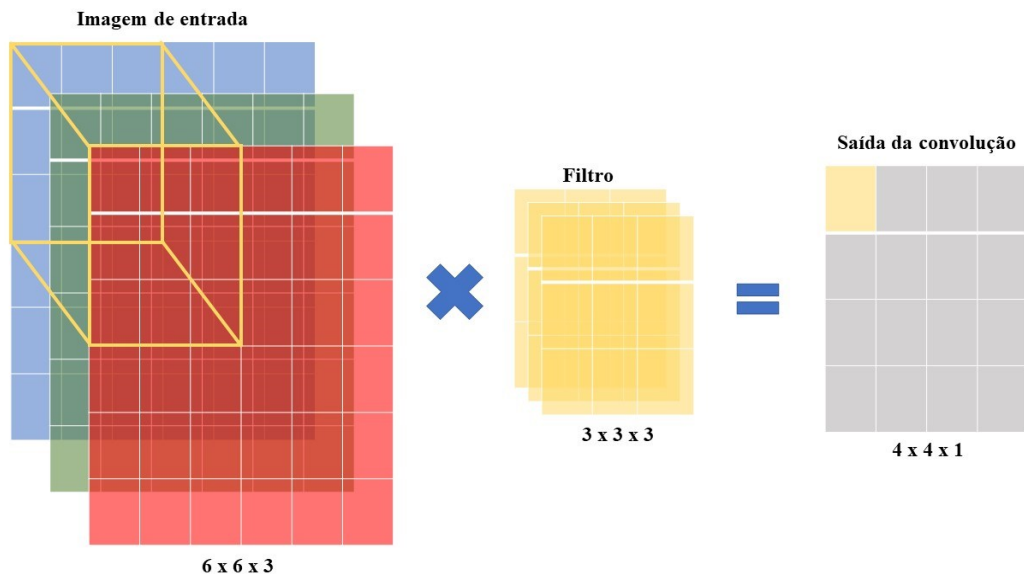
Figura 20 – Representação de tensores unidimensional 1x6, bidimensional 6x6 e tridimensional 6x6x3.



Fonte: Autoria própria (2020)

Para demonstrar como funciona os filtros em camadas convolucionais vamos utilizar como exemplo o tensor tridimensional da figura 20, onde vamos aplicar um filtro 3x3x3 e demonstrar o resultado obtido. A figura 21 demonstra a utilização de um filtro convolucional em uma imagem de entrada de tamanho 6x6 pixels e 3 camadas de cores RGB.

Figura 21 – Representação de aplicação de filtro 3x3 em uma imagem 6x6 com 3 canais de cores e seu resultado.



Fonte: Autoria própria (2020)

O sistema prevê caixas delimitadoras usando grupo de dimensões como caixas de ancoragem, por exemplo, prevê 4 coordenadas para cada caixa delimitadora. São calculados erros com base no deslocamento da célula em relação às coordenadas, como é representado na figura 22. Como exemplo, para as coordenadas t_x, t_y, t_w, t_h quando a célula estiver deslocada do canto superior esquerdo da imagem por (c_x, c_y) , a caixa anterior possui largura e altura p_w, p_h as previsões correspondem as equações 12 a 15 (REDMON e FARHADI, 2018).

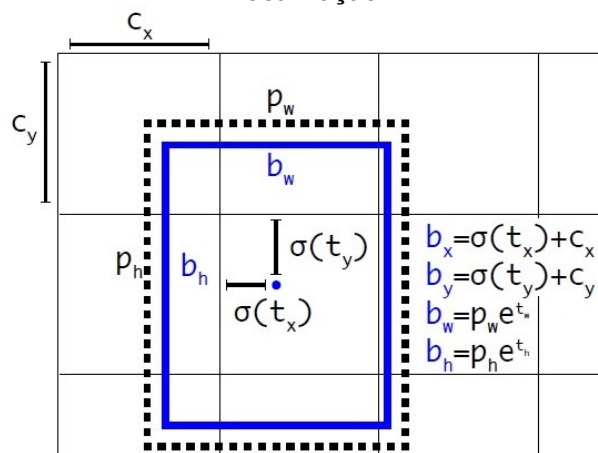
$$b_x = \sigma(t_x) + c_x \quad (12)$$

$$b_y = \sigma(t_y) + c_y \quad (13)$$

$$b_w = p_w e^{t_w} \quad (14)$$

$$b_h = p_h e^{t_h} \quad (15)$$

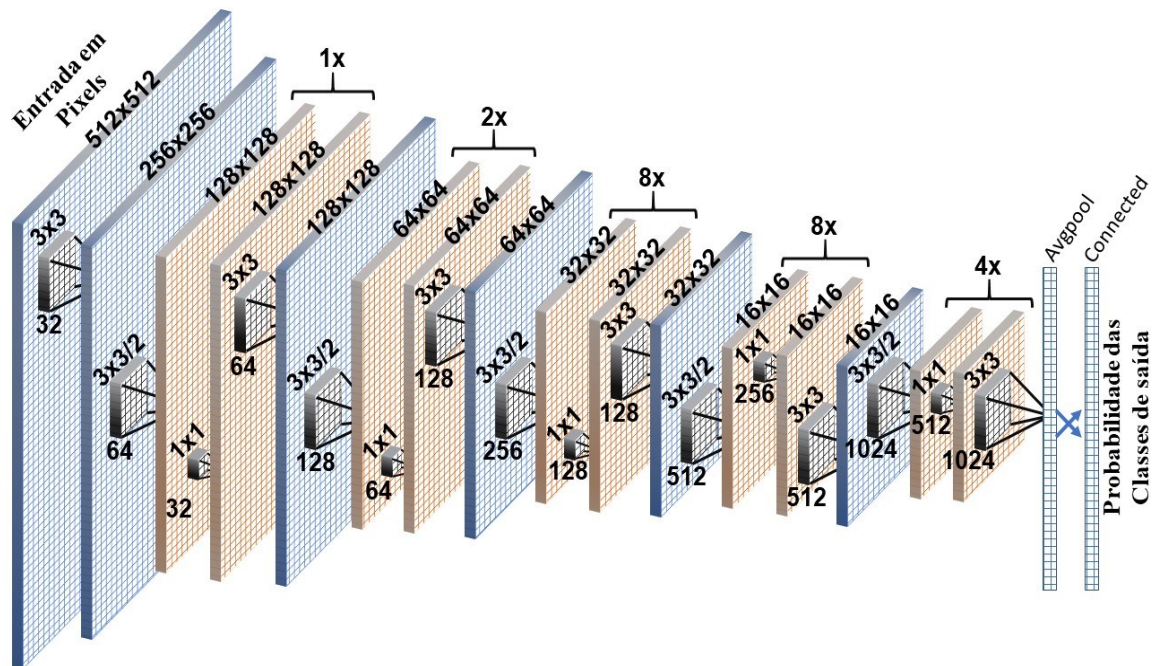
Figura 22 – Representação das caixas delimitadoras com dimensões anteriores e previsão de localização.



Fonte: Adaptado de REDMON e FARHADI (2018)

A versão YOLO que estamos utilizando nesta pesquisa é baseado em uma rede DarkNet de 53 camadas convolucionais, chamada Darknet-53. Esta é uma rede mais poderosa que a utilizada nas versões anteriores com 19 camadas (REDMON e FARHADI 2018) e mais eficiente que ResNet-101 ou ResNet-152 (REDMON e FARHADI 2018). A figura 23 representa como a rede é distribuída, demonstra de forma visual como a rede se comporta no momento do treinamento com relação ao redimensionamento da entrada da rede, a aplicação dos filtros até a saída.

Figura 23 – Representação da distribuição das camadas na rede YOLOv3.



Fonte: Autoria própria (2020)

A rede neural conta com alguns hiperparâmetros, a entrada da rede pode ser configurada utilizando qualquer valor múltiplo de 32 (e.g., 32x32, 64x64, ..., 512x512), nos nossos treinamentos utilizamos o valor de entrada de 512x512, a imagem passa pelas camadas convolucionais, estas camadas funcionam como filtros, passando por todas as camadas da imagem e extraíndo suas características, formando um mapa de características. Os filtros ou kernel, como são conhecidos são formados por pesos de valores aleatórios, que se atualizam a cada nova entrada no processo de backpropagation. Complementando o filtro ainda temos o *stride* e o *padding*. O *stride*, corresponde ao passo do nosso filtro que no YOLOV3 varia entre os valores 1 e 2. O *padding* tem como principal função fazer com que o filtro seja aplicado o mesmo número de vezes para todos os pixels, a saída da convolução ficará do mesmo tamanho ou zero *padding*, onde uma borda com zeros é adicionada.

O lote (*batch*) que representa quantas imagens serão processadas ao mesmo tempo, subdivisões (*subdivisions*) que representa quantas imagens do lote serão analisadas por vez, estes parâmetros influenciam na velocidade do treino por conta do processamento em lotes com respectiva atualização de pesos da rede neural. O treinamento se torna mais rápido com o processamento em lotes maiores, porém os parâmetros precisaram ser ajustados de acordo com a capacidade de processamento do hardware utilizado. Para os experimentos realizados foi utilizado o valor de *batch*

igual a 20 e *subdivisions* igual a 5. Neste trabalho utilizou-se um processador Intel Core I7-9750H CPU @ 2.60Hz x 12, 16GB de RAM, 512GB de SSD e placa de Vídeo Nvidia GeForce RTX 2060 com 6GB de VRAM GDDR6 e 1920 CUDA cores.

Largura e altura (*width* e *height*) que representam os valores de largura e altura em pixels, para os quais as imagens de entrada serão redimensionadas na entrada da rede neural.

O canal (*channel*) determina o número de canais de entrada da imagem que será passada para a rede neural, para este parâmetro utilizamos o valor 3 devido o item de entrada ser uma imagem que pode ser comparada a uma matriz de 3 dimensões onde cada dimensão representa um canal RGB.

Momento (*momentum*) e decaimento (*decay*) que são parâmetros da função gradiente descendente, na qual o momento define o peso do gradiente previamente calculado no novo gradiente e o decaimento e a taxa de decaimento do aprendizado. Para esses parâmetros foram usados os valores de 0.9 e 0,0005, respectivamente.

Os parâmetros ângulo (*angle*), saturação (*saturation*), exposição (*exposure*) a luz e tonalidade (*hue*), que são parâmetros que definem o aumento dos dados (Data Augmentation), imagens adicionais auxiliares para o treinamento. Foram utilizados os valores 0 para ângulo (i.e., variação aleatória em ampla faixa de ângulos) e 1,5, 1,5 e 0,1 para saturação, exposição a luz e tonalidade, respectivamente, que representam limites, para mais ou para menos, dentro dos quais as imagens podem ser aleatoriamente modificadas, a tabela 1 mostra os hiperparâmetros configurados para os experimentos A, B, C, D, E e F realizados no YOLO.

Tabela 1 – Hiperparâmetros configurados na para o treinamento, validação e análise de desempenho.

Hiperparâmetros	Valor
Total de imagens utilizadas para treinamento e análise de erros	Variável por experimento
Treinamento	80%
Análise de erros	20%
<i>Batch</i>	16
<i>Subdivision</i>	8
<i>Width X Heigth</i>	512 x 512
<i>Chanel</i>	3
<i>Momentum</i>	0.9
<i>Decay</i>	5
<i>Angle</i>	0
<i>Saturation</i>	1.5
<i>Exposure</i>	1.5
<i>Hue</i>	0.1

Fonte: Autoria própria (2020)

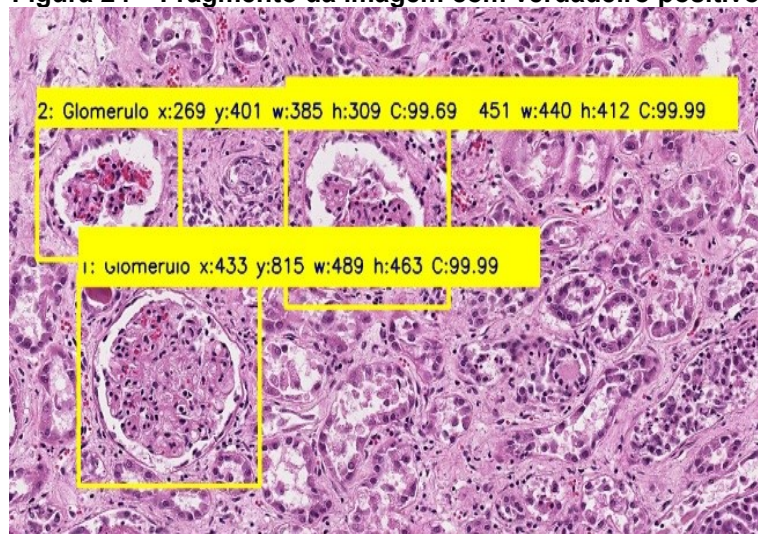
Função de ativação, traz a não linearidade para o sistema, de forma que a rede possa aprender qualquer tipo de funcionalidade. Há diversas funções como a Sigmoid, TanH, ReLU, *Leaky ReLU* entre outras, no YOLOv3 a que é utilizada é a *Leaky ReLU*.

Camadas de Pooling e a full connected: a camada de pooling, no caso do YOLOv3 é a camada *avgpool (Average Pooling)*, é uma camada de *down-sampling*, tem como função simplificar as informações das camadas anteriores, reduzindo o tamanho espacial das matrizes resultantes das convoluções, o que contribui para o controle de overfitting. A camada full connected é necessária para converter a matriz resultante da camada de pooling em um vetor com as características encontradas no aprendizado.

2.6 Métricas para avaliação de desempenho

As métricas são calculadas baseadas nos parâmetros: verdadeiro positivo (i.e., TP, *true positive*), falso positivo (i.e., FP, *false positive*), falso negativo (i.e., FN, *false negative*) e verdadeiro negativo (i.e., TN, *true negative*). As figuras 24 a 26 demonstram o entendimento destes parâmetros (i.e., TP, FP, FN e TN) utilizando imagens para exemplificar. TP é uma região com glomérulo (i.e., Positivo) que é marcada como glomérulo como apresentado na figura 24.

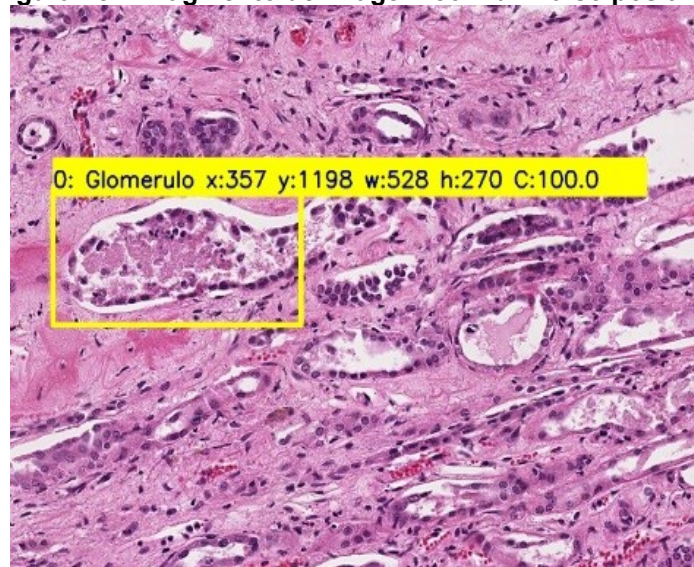
Figura 24 – Fragmento da imagem com verdadeiro positivo.



Fonte: Autoria própria (2020)

FP é uma região sem glomérulo que é marcada como sendo glomérulo. Um fragmento de imagem representando este parâmetro é apresentado na figura 25.

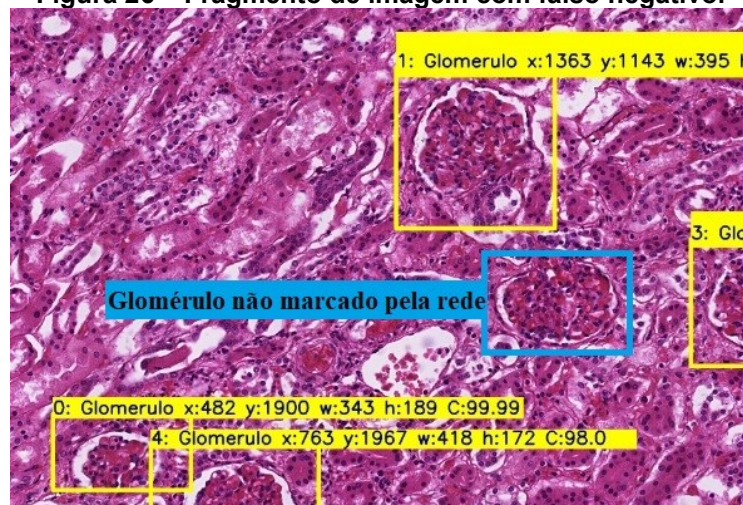
Figura 25 – Fragmento de imagem com um falso positivo.



Fonte: Autoria própria (2020)

FN é uma região com glomérulo que é reconhecida como não tendo glomérulo. Isto é representado na região delimitada por um retângulo azul na figura 26.

Figura 26 – Fragmento de imagem com falso negativo.

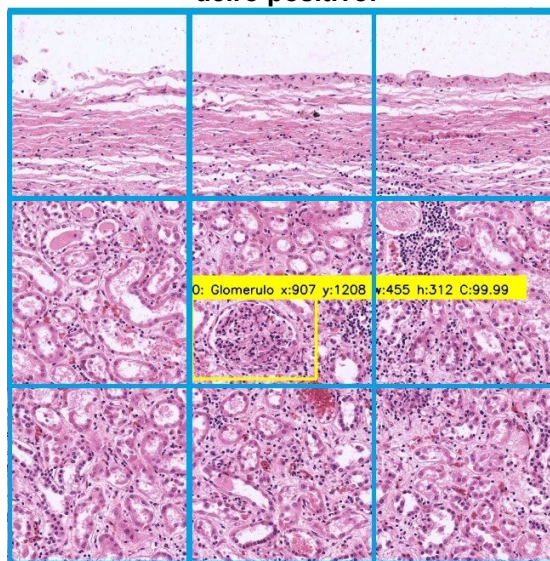


Fonte: Autoria própria (2020)

TN é uma região sem glomérulo (i.e., Negativo) que não é marcada como tendo glomérulo, ou seja, região verdadeiramente reconhecida como não tendo glomérulo. Este é o caso da maioria dos pixels de uma imagem para os casos utilizados nesta pesquisa. Desta forma, inicialmente computou-se \bar{W} e \bar{H} que correspondem às médias da altura e largura dos glomérulos marcados nas imagens. Da mesma forma, computou-se σ_W e σ_H que correspondem aos desvios padrão da altura e largura dos glomérulos marcados nas imagens. Considera-se que o tamanho dos glomérulos tem

uma distribuição Gaussiana. Desta forma, pode-se esperar que pelo menos 95% dos glomérulos que podem ser identificados caberiam em uma região definida como tendo a área de $(\bar{H} + 4 * \sigma_H) * (\bar{W} + 4 * \sigma_W)$. Portanto, o tamanho típico dos retângulos azuis marcados como não tendo glomérulo encontrado é dado por esta equação. Cada imagem apresentada à rede neural, possui o tamanho dado por $(W * H)$. Desta forma, através da equação 16, pode-se inferir a quantidade de regiões em que caberiam glomérulos típicos e que os mesmos não foram identificados. Isto representa neste trabalho a quantidade de TN. A figura 27 representa como é efetuada a contagem dos espaços sem marcação de glomérulos em uma imagem permitindo que regiões onde tipicamente caberiam glomérulos, mas que para as quais os mesmos não foram identificados, sejam contabilizadas como TN.

Figura 27 – Representação das áreas calculadas através da aplicação da equação 16 para contagem dos verdadeiros negativos. Nesta imagem existem 8 verdadeiros negativos e um verdadeiro positivo.



Fonte: Autoria própria (2020)

Para a avaliação dos resultados obtidos utilizou-se como métricas: Precisão (P) (i.e., valor de predições positivas) que representa a possibilidade de uma marcação feita estar correta; Sensibilidade (S) que é a probabilidade de um glomérulo ser encontrado (KAWAMURA, 2002); Acurácia (A) que determina o quão próximo os resultados obtidos estão dos resultados reais (EUSEBI, 2013); Especificidade (E) que determina a probabilidade de uma região não detectada não conter um glomérulo marcado erroneamente (i.e., FP); VPN (i.e., valor de predições negativas) que representa a probabilidade de áreas não detectadas não conterem um glomérulo não

identificado (i.e., F.N.) (KAWAMURA, 2002) e F1-Score (F1s) que representa a média harmônica entre a precisão e a sensibilidade (CHICCO & JURMAN, 2020). As equações 17 a 22 foram utilizadas para determinar, precisão, sensibilidade, acurácia, especificidade, VPN e F1s, respectivamente. Para os cálculos foram considerados os TP, FN, FP e TN.

$$Tn = \frac{(W * H) - N_G * ((\bar{H} + 4 * \sigma_H) * (\bar{W} + 4 * \sigma_W))}{((\bar{H} + 4 * \sigma_H) * (\bar{W} + 4 * \sigma_W))} \quad (16)$$

$$P = \frac{TP}{TP + FP} \quad (17)$$

$$S = \frac{TP}{TP + FN} \quad (18)$$

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

$$E = \frac{TN}{TN + FP} \quad (20)$$

$$VPN = \frac{TN}{TN + FN} \quad (21)$$

$$F1_s = \frac{2 * P * S}{P + S} \quad (22)$$

3 MATERIAIS E MÉTODOS

Este capítulo descreve o método de pesquisa utilizado para o desenvolvimento do projeto de aprendizagem profunda aplicada à análise de histopatologia renal usando o framework YOLOv3.

3.1 Base de dados

As imagens histopatológicas (i.e., total de 16 lâminas) foram obtidas de banco de imagens de patologia citadas pela organização internacional de Patologia Digital. Utilizamos as bases Juan Rosai's, Bio-Atlas e University of Leeds, de acesso livre às lâminas de patologia contendo imagens de cortes histológicos renais, sem necessidade de login e senha para acesso.

3.1.1 Base de dados

Base de dados disponibilizada pelo Dr. Juan Rosai em parceria com a *USCAP United State & Canadian Academy of Pathology* e a Aperio. Esta coleção possui uma base de aproximadamente 20.000 casos referentes a seminários de patologias desde o ano de 1945 até o presente. Esta coleção inclui resumos de históricos clínicos, discussões e diagnósticos realizados pelo Dr. Rosai e vários outros especialistas (ROSAI n.d., 1945).

O acesso das imagens é através do site <http://rosai.secondslide.com/>, onde é possível visualizar e efetuar o download das imagens com a auxílio do software ImageScope.

Da base de dados Rosai's foram utilizadas 7 lâminas do seminário 131, casos 8, 10, 11, 12, 13, 14 e 15, renomeadas para Case8_F, Case10_G, Case11_H, Case12_I, Case13_J, Case14_K e Case15_L respectivamente, estas imagens foram salvas local com um zoom de 20x, o que proporciona uma melhor visualização das estruturas utilizadas para todo o processo (i.e., treinamento, validação). Estas imagens foram utilizadas para o treinamento dos experimentos A, B, C, D, E e F.

3.1.2 Base Bio-Atlas

Base de dados disponibilizada por Jake Gittlen *Laboratories for Cancer Research – The Pennsylvania State University*. Esta coleção foi originalmente montada para peixes-zebras e tecidos humanos, atualmente possui diversos parceiros que

auxiliam para ampliar a base mantendo-a atualizada e diversificada (FOUNDATION e FUND, 2007).

O acesso das imagens é através do site <https://bio-atlas.psu.edu/human/index.php>, onde é possível visualizar e efetuar o download das imagens.

Da base de dados Bio-Atlas foram utilizadas 5 lâminas, as lâminas 837, 1060, 1267, 1292 e 1298, renomeadas para *Img_837*, *Img_1060*, *Img_1267*, *Img_1292* e *Img_1298*, respectivamente. Estas imagens foram salvas no computador local com um zoom de 20x, proporcionando uma melhor visualização das estruturas utilizadas para todo o processo (i.e., treinamento, validação). Estas imagens foram utilizadas para o treinamento dos experimentos B, C e D.

As imagens utilizadas na avaliação de desempenho da rede (i.e., lâminas desconhecidas da rede YOLO, não utilizadas para o treinamento) foram utilizadas 5 lâminas, são lâminas da base de dados Bio-Atlas, porém, são outras lâminas (i.e., lâminas desconhecidas da rede YOLO), são elas as lâminas 569, 603, 651, 849 e 1307, renomeadas para *Img569*, *Img603*, *Img651*, *Img849* e *Img1307*, respectivamente. Estas imagens foram salvas no computador local com um zoom de 20x, e foram utilizadas somente para a avaliação de desempenho da rede.

3.1.3 Base University of Leeds

Base de dados disponibilizada pela *University of Leeds* que possui uma equipe dedicada para digitalização e hospedagem de slides digitais em alta qualidade atualmente com aproximadamente 150TB de dados de slides digitais. É um site dedicado para fornecer recursos para o ensino, pesquisa e trabalho clínico de patologia digital (Dept 2003).

O acesso das imagens é através do site <https://www.virtualpathology.leeds.ac.uk/slides/library/>, onde é possível visualizar e efetuar o download das imagens com a auxílio do software ImageScope.

Da base de dados Leeds foram utilizadas 4 lâminas, as lâminas 1057 e 33814, 134149 e 149335 renomeadas para *1057_A* e *33814_B*, *134149_D* e *149335_E* respectivamente. Estas imagens foram salvas no computador local com um zoom de 20x, proporcionando uma melhor visualização das estruturas utilizadas para todo o processo (i.e., treinamento, validação). Estas imagens foram utilizadas para o treinamento dos experimentos A, C, D, E e F.

As imagens utilizadas na avaliação de desempenho são as lâminas 1111, 33836 e 43267, renomeadas para lmg1111, lmg33836 e lmg43267 respectivamente. Estas imagens foram salvas no computador local com um zoom de 20x, e foram utilizadas somente para a avaliação de desempenho da rede.

3.2 Preparo das imagens

Para a utilização das imagens nos treinamentos, validação da rede e avaliação de desempenho, foi necessário ter os arquivos digitais no computador em que a rede e o algoritmo de treinamento estão instalados. No caso dos repositórios Leeds e Rosai's, o download das imagens foi simples, não foi necessário comprar ou desenvolver algum algoritmo para baixar as imagens, só a instalação do software Aperio ImageScope, software gratuito que pode ser baixado através do site da Leica (i.e., <http://www2.leicabiosystems.com/l/48532/2019-01-03/7bs4rg>). No momento do desenvolvimento deste documento, a versão utilizada é a 12.4.3. Porém, de acordo com o andamento do projeto, tivemos alguns problemas devido ao tamanho das imagens pois ambas as imagens utilizadas têm uma resolução de aproximadamente 30000 a 50000 pixels de altura por 20000 a 40000 pixels de largura. O processamento de imagens com esta quantidade de pixels ainda é desafiador para os sistemas computacionais e softwares vigentes.

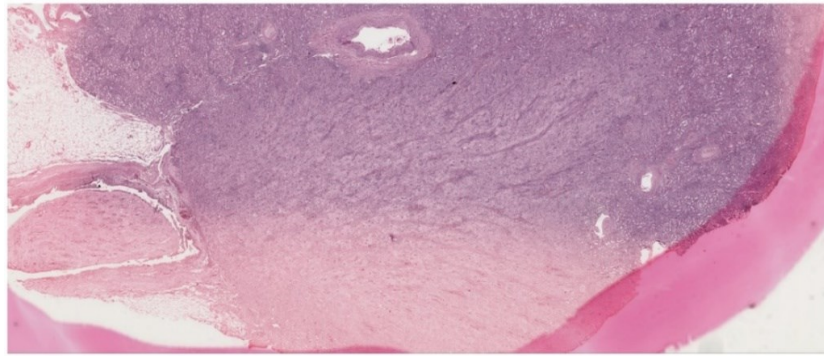
Após a escolha da imagem baixamos o arquivo com a extensão específica (e.g., .sis) para a abertura pelo software Aperio, este software possibilitou o download das imagens para máquina local. Devido ao tamanho das imagens, foi necessário dividi-las em duas partes, uma imagem como parte superior e outra como parte inferior. A figura 28 demonstra as duas partes da imagem Case10_G, após o download.

Figura 28 – Representação do download da imagem Case10_G separada em duas partes, a imagem (a) representa a parte superior da imagem e a imagem (b) a parte inferior ambas com tamanho de 37980x13698.

(a)



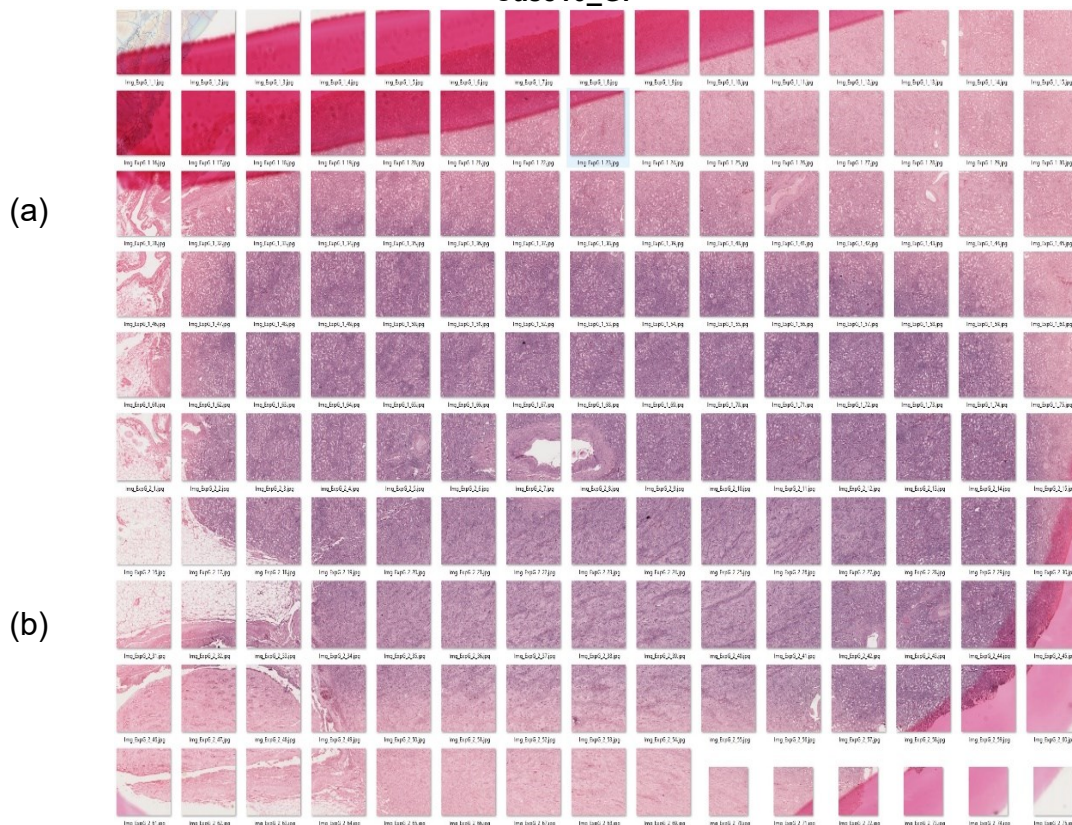
(b)



Fonte: Autoria própria (2020)

Mesmo dividindo a imagem em duas partes tivemos problemas de processamento (e.g., estouro de memória). Para contornar os erros, as imagens foram recortadas. Para cada imagem foi padronizada a resolução de 2048x2048. A figura 29 representa as partes superior e inferior da imagem Case10_G recortada no tamanho padronizado de 2048x2048, (a) representa a parte superior e (b) representa a parte inferior da imagem Case10_G.

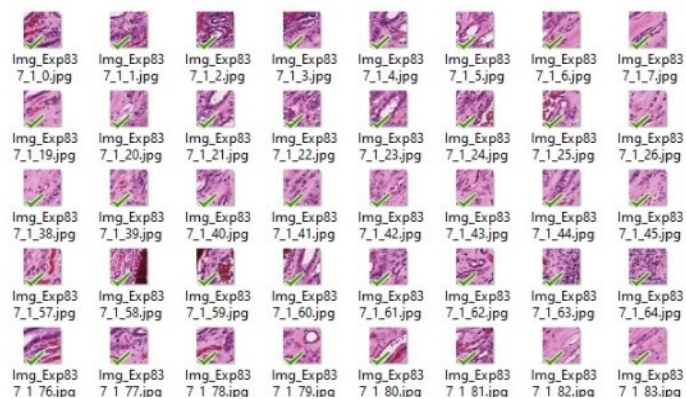
Figura 29 – Representação da imagem Case10_G após o recorte para o tamanho padrão de 2048x2048, (a) representa a parte superior da imagem Case10_G, (b) representa a parte inferior Case10_G.



Fonte: Autoria própria (2020)

As imagens disponibilizadas no site Bio-atlas são montadas a partir de várias imagens pequenas, com resolução de 256x256 pixels cada. O download é realizado e 12x12 imagens de 256x256 pixels são justapostas para formar as imagens na resolução padrão definida para a rede (i.e., 2048x2048). Na figura 30 podemos ver as 5.168 imagens com resolução de 256x256 que formam na sequência 837 imagens com tamanho padronizado de 2048x2048.

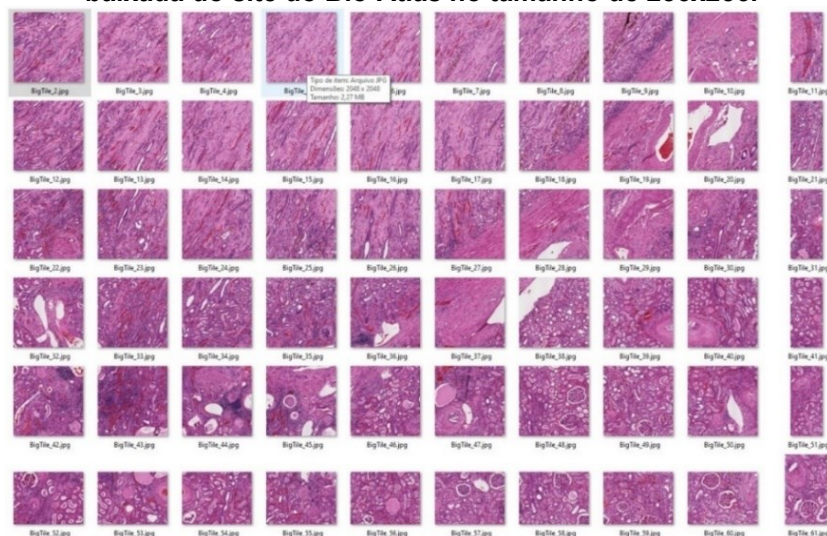
Figura 30 – Exemplificação dos 5.168 fragmentos de imagens de tamanho 256x256 referente a imagem 837 retiradas do site do Bio-Atlas.



Fonte: Autoria própria (2020)

Com as imagens armazenadas localmente utilizou-se o segundo desenvolvimento que monta as imagens no tamanho padrão (i.e., 2048x2048), a partir das imagens de 256x256. A figura 31 mostra a representação das imagens montadas no tamanho 2048x2048.

Figura 31 – Representação das 60 imagens de 2048x2048 formadas a partir das imagens 837 baixada do site do Bio-Atlas no tamanho de 256x256.

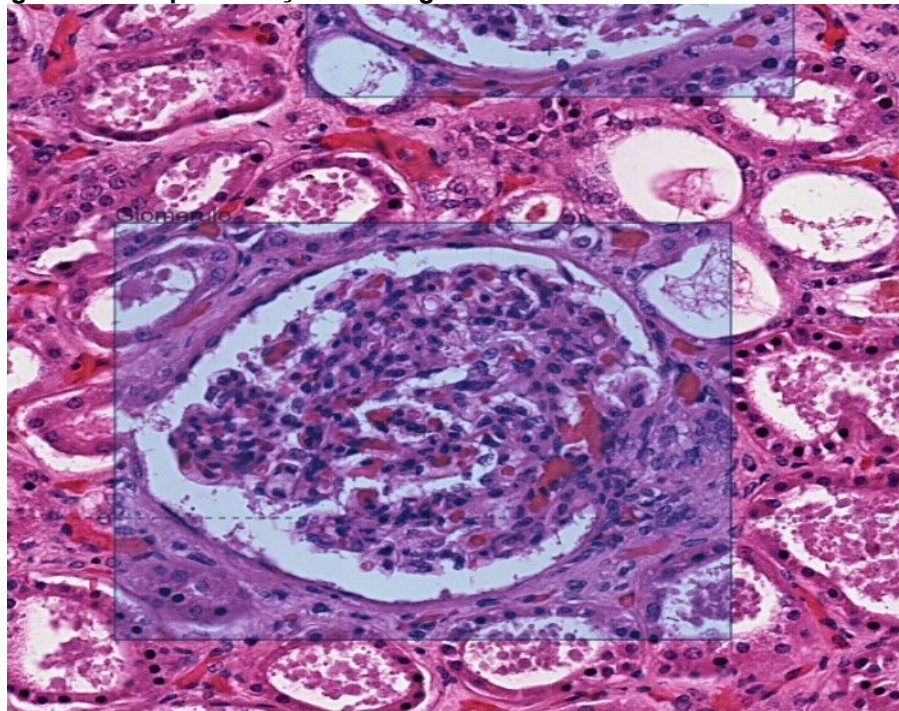


Fonte: Autoria própria (2020)

3.2.1 Anotação das imagens

Como a rede utiliza o aprendizado supervisionado (i.e., a rede necessita saber a estrutura e a classe do objeto), foi necessário realizar a anotação das imagens. Este processo efetuado de forma manual com o auxílio do software YBAT (YOLO BBox Annotation Tool), desenvolvido em Java Script e com uma interface gráfica amigável para anotação das imagens (SUN, 2020). Este processo foi possível a partir do conhecimento de médicos residentes de patologista que demonstraram como é efetuada a identificação das estruturas glomerulares em lâminas de histopatologia e o que é o glomérulo. A figura 32 mostra a imagem anotada através do software YBAT.

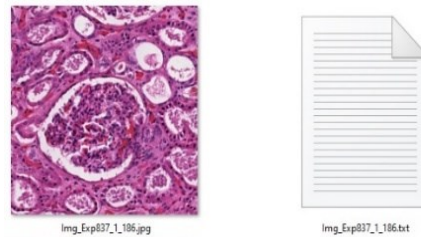
Figura 32 – Representação da imagem anotada através do software YBAT.



Fonte: Autoria própria (2020)

Para cada imagem anotada é gerado um arquivo texto com as informações de localização e eventualmente classes dos objetos procurados (i.e., glomérulo), coordenadas (i.e., c_x e c_y), altura e largura (i.e., p_w e p_h), dos itens marcados nas imagens, a figura 33 representa a imagem e o arquivo gerado após a anotação com as informações de classe, coordenadas, altura e largura (e.g., 0 0.4638671875 0.60302734375 0.68359375 0.5869140625) para o objeto (i.e., glomérulo), marcado no centro da imagem.

Figura 33 – Representação da imagem e do arquivo gerado pelo programa YBAT após a anotação da imagem.



Fonte: Autoria própria (2020)

Como citado anteriormente, para esta pesquisa foram utilizadas 16 lâminas de imagens inteiras (i.e., área total de uma lâmina) que foram subdivididas, formando 815 imagens com resolução de 2048x2048 onde foram marcados 2325 glomérulos. Destas imagens, 80% foram separadas para o treinamento e 20% foram separadas para validação (i.e., imagens utilizadas pela própria rede para avaliar o treinamento da rede). A tabela 2 demonstra as lâminas utilizadas (i.e., nome dos arquivos), repositórios (i.e., sites de onde foram efetuados os downloads das imagens), o total de imagens (i.e., resolução de 2048x2048) utilizadas e o total de glomérulos anotados por imagens.

Tabela 2 – Demonstra as lâminas utilizadas, os repositórios utilizados para download das imagens número total de imagens de 2048x2048 e o total de glomérulos anotados por imagens.

	Repositórios	Total de imagens (Qtd)	Total de Glomérulos (Qtd)
Lâminas
1057_A	Leeds	101	246
33814_B	Leeds	83	109
1341159_D	Leeds	20	37
149335_E	Leeds	78	154
Imagem_837	Bio-Atlas	25	47
Imagem_1060	Bio-Atlas	32	88
Imagem_1267	Bio-Atlas	15	33
Imagem_1292	Bio-Atlas	43	115
Imagem_1298	Bio-Atlas	36	76
Case8_F	Rosai's	48	191
Case10_G	Rosai's	61	186
Case11_H	Rosai's	77	279
Case12_I	Rosai's	30	151
Case13_J	Rosai's	72	312
Case14_K	Rosai's	62	165
Case15_L	Rosai's	32	136
Total	3	815	2325

Fonte: Autoria própria (2020)

3.3 Treinamento

Para o treinamento, as imagens utilizadas para a base de imagens foram separadas em 6 experimentos denominados experimentos A, B, C e D, onde o A utiliza

as bases de imagens criada a partir das imagens retiradas dos sites Leeds e Rosai's conforme apresentado na Tabela 3, B utiliza as bases de imagens criadas com das imagens retiradas dos sites Bio-Atlas e Rosai's conforme apresentado na Tabela 4, C utiliza as mesmas imagens contidas em A e B conforme apresentado na Tabela 5, D utiliza as mesmas imagens contidas em C, com a base duplicada conforme apresentado na Tabela 6, para o treinamento de A a D o tempo de processamento foi de aproximadamente 20 horas, 50000 iterações. Os experimentos E e F, onde E utiliza as mesmas imagens apresentadas na Tabela 3, e F utilizou as mesmas imagens da base de imagens da Tabela 3 duplicadas, como podemos ver na Tabela 7. Os hiperparâmetros foram modificados para permitir um treinamento com 60000 interações, o que aumentou o tempo de processamento do treinamento em aproximadamente 6 horas. Cada experimento foi separado em dois grupos de imagens, sendo eles, o grupo de treinamento da rede contendo 80% das imagens e o grupo de teste da rede contendo 20% das imagens, o conjunto de teste da rede é utilizado para determinar a performance da rede, são imagens que não foram previamente utilizados no momento do treinamento, que a rede utiliza para calcular o quanto está aprendendo. O motivo utilizado para a separação dos datasets em experimentos (e.g., A, B, C, D, E e F), foi analisar e identificar qual a melhor resposta entre os treinamentos efetuados, através do progresso obtido com a diversificação e duplicação das imagens entre os datasets de treinamentos. As tabelas 3 a 7 mostram as lâminas utilizadas, o repositório de onde foram extraídas cada lâmina, o total de imagens geradas para cada lâmina, o total de glomérulos marcado na fase de anotação, o total de imagens utilizada para o treinamento com seus respectivos números de glomérulos e as imagens de teste da rede também com os seus respectivos números de glomérulos, e como ficaram separados os datasets de cada experimento para os treinamentos.

Tabela 3 – Tabela com informações da base de imagens utilizadas para o treinamento e validação dos experimentos A e E.

(continua)

	Rep.	Img. (Qtd)	Glom. (Qtd)	Img. treino (Qtd)	Glom. treino (Qtd)	Img. Valid. (Qtd)	Glom. Valid.(Qtd)
Lâminas
1057_A	Leeds	101	246	85	210	16	36
33814_B	Leeds	83	109	70	93	13	16
1341159_D	Leeds	20	37	13	21	7	16
149335_E	Leeds	78	154	62	123	16	31

Tabela 3 – Tabela com informações da base de imagens utilizadas para o treinamento e validação dos experimentos A e E.

(conclusão)

	Rep.	Img. (Qtd)	Glom. (Qtd)	Img. treino (Qtd)	Glom. treino (Qtd)	Img. Valid. (Qtd)	Glom. Valid.(Qtd)
Lâminas
Case8_F	Rosai's	48	191	36	144	12	47
Case10_G	Rosai's	61	186	53	154	8	32
Case11_H	Rosai's	77	279	62	227	15	52
Case12_I	Rosai's	30	151	24	130	6	21
Case13_J	Rosai's	72	312	60	260	12	52
Case14_K	Rosai's	62	165	44	121	18	44
Case15_L	Rosai's	32	136	23	99	9	37
Total	2	664	1966	532	1582	132	384

Fonte: Autoria própria (2020)

Tabela 4 – Tabela com informações da base de imagens utilizadas para o treinamento e validação do experimento B.

	Rep.	Img. (Qtd)	Glom. (Qtd)	Img. treino (Qtd)	Glom. treino (Qtd)	Img. Valid. (Qtd)	Glom. Valid.(Qtd)
Lâminas
Img_837	Bio-Atlas	25	47	21	39	4	8
Img_1060	Bio-Atlas	32	88	24	65	8	23
Img_1267	Bio-Atlas	15	33	11	25	4	8
Img_1292	Bio-Atlas	43	115	35	92	8	23
Img_1298	Bio-Atlas	36	76	27	59	9	17
Case8_F	Rosai's	48	191	36	144	12	47
Case10_G	Rosai's	61	186	53	154	8	32
Case11_H	Rosai's	77	279	62	227	15	52
Case12_I	Rosai's	30	151	24	130	6	21
Case13_J	Rosai's	72	312	60	260	12	52
Case14_K	Rosai's	62	165	44	121	18	44
Case15_L	Rosai's	32	136	23	99	9	37
Total	2	533	1779	420	1415	113	364

Fonte: Autoria própria (2020)

Tabela 5 – Tabela com informações da base de imagens utilizadas para o treinamento e validação do experimento C.

(continua)

	Rep.	Img. (Qtd)	Glom. (Qtd)	Img. treino (Qtd)	Glom. treino (Qtd)	Img. Valid. (Qtd)	Glom. Valid.(Qtd)
Lâminas
1057_A	Leeds	101	246	85	210	16	36
33814_B	Leeds	83	109	70	93	13	16
134159_D	Leeds	20	37	13	21	7	16
149335_E	Leeds	78	154	62	123	16	31
Img_837	Bio-Atlas	25	47	21	39	4	8
Img_1060	Bio-Atlas	32	88	24	65	8	23
Img_1267	Bio-Atlas	15	33	11	25	4	8
Img_1292	Bio-Atlas	43	115	35	92	8	23
Img_1298	Bio-Atlas	36	76	27	59	9	17
case8_F	Rosai's	48	191	36	144	12	47
case10_G	Rosai's	61	186	53	154	8	32
case11_H	Rosai's	77	279	62	227	15	52

Tabela 5 – Tabela com informações da base de imagens utilizadas para o treinamento e validação do experimento C.

(conclusão)

	Rep.	Img. (Qtd)	Glom. (Qtd)	Img. treino (Qtd)	Glom. treino (Qtd)	Img. Valid. (Qtd)	Glom. Valid.(Qtd)
Lâminas
case12_I	Rosai's	30	151	24	130	6	21
case13_J	Rosai's	72	312	60	260	12	52
case14_K	Rosai's	62	165	44	121	18	44
case15_L	Rosai's	32	136	23	99	9	37
Total	3	815	2325	650	1862	165	463

Fonte: Autoria própria (2020)

Tabela 6 – Tabela com informações da base de imagens utilizadas para o treinamento e validação do experimento D.

	Rep.	Img. (Qtd)	Glom. (Qtd)	Img. treino (Qtd)	Glom. treino (Qtd)	Img. Valid. (Qtd)	Glom. Valid.(Qtd)
Lâminas
1057_A	Leeds	202	492	170	420	32	72
33814_B	Leeds	166	218	140	186	26	32
134159_D	Leeds	40	74	26	42	14	32
149335_E	Leeds	156	308	124	246	32	62
Img_837	Bio-Atlas	50	94	42	78	8	16
Img_1060	Bio-Atlas	64	176	48	130	16	46
Img_1267	Bio-Atlas	30	66	22	50	8	16
Img_1292	Bio-Atlas	86	230	70	184	16	46
Img_1298	Bio-Atlas	72	152	54	118	18	34
Case8_F	Rosai's	96	382	72	288	24	94
Case10_G	Rosai's	122	372	106	308	16	64
Case11_H	Rosai's	154	558	124	454	30	104
Case12_I	Rosai's	60	302	48	260	12	42
Case13_J	Rosai's	144	624	120	520	24	104
Case14_K	Rosai's	124	330	88	242	36	88
Case15_L	Rosai's	64	272	46	198	18	74
Total	3	1630	4650	1300	3724	330	926

Fonte: Autoria própria (2020)

Tabela 7 – Tabela com informações da base de imagens utilizadas para o treinamento e validação do experimento F.

(continua)

	Rep.	Img. (Qtd)	Glom. (Qtd)	Img. treino (Qtd)	Glom. treino (Qtd)	Img. Valid. (Qtd)	Glom. Valid.(Qtd)
Lâminas
1057_A	Leeds	202	492	170	420	32	72
33814_B	Leeds	166	218	140	186	26	32
134159_D	Leeds	40	74	26	42	14	32
149335_E	Leeds	156	308	124	246	32	62
Case8_F	Rosai's	96	382	72	288	24	94
Case10_G	Rosai's	122	372	106	308	16	64
Case11_H	Rosai's	154	558	124	454	30	104
Case12_I	Rosai's	60	302	48	260	12	42
Case13_J	Rosai's	144	624	120	520	24	104

Tabela 7 – Tabela com informações da base de imagens utilizadas para o treinamento e validação do experimento F.

(conclusão)

	Rep.	Img. (Qtd)	Glom. (Qtd)	Img. treino (Qtd)	Glom. treino (Qtd)	Img. Valid. (Qtd)	Glom. Valid.(Qtd)
Lâminas
Case14_K	Rosai's	124	330	88	242	36	88
Case15_L	Rosai's	64	272	46	198	18	74
Total	2	1328	3932	1064	3164	264	768

Fonte: Autoria própria (2020)

3.4 Avaliação de desempenho

A avaliação de desempenho do sistema treinado, foi efetuada com 7 lâminas de biópsia contendo glomérulos, as 7 lâminas são desconhecidas (i.e., não fizeram parte da base de dados utilizada para o treinamento) da rede, têm diferentes tonalidades e diferentes proporções de glomérulos. As tabelas 8 e 9 mostram a quantidade de imagens e glomérulos contidos nestas lâminas.

Tabela 8 – Imagens do repositório Bio-Atlas utilizadas para avaliação de desempenho da rede.

	Repositório.	Img. (Qtd)	Glom. (Qtd)
Lâminas
Img_569	Bio-Atlas	60	116
Img_1307	Bio-Atlas	144	152
Img_603	Bio-Atlas	81	33
Img_651	Bio-Atlas	80	41
Img_849	Bio-Atlas	60	123
Total	1	425	465

Fonte: Autoria própria (2020)

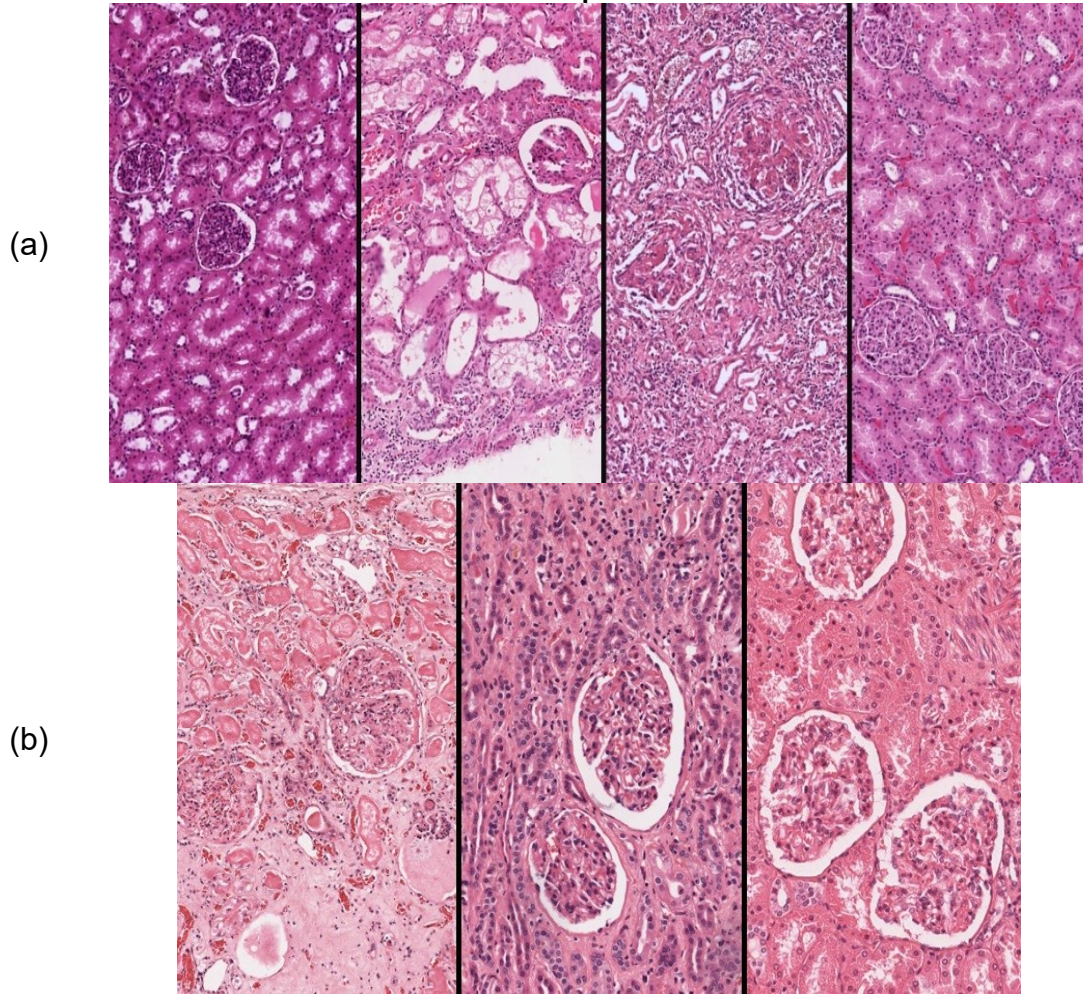
Tabela 9 – Imagens do repositório Leeds utilizadas para avaliação de desempenho da rede.

	Repositório.	Img. (Qtd)	Glom. (Qtd)
Lâminas
Imgem_1111	Leeds	592	290
Imgem_33836	Leeds	1014	407
Imgem_43267	Leeds	768	242
Total	1	2374	939

Fonte: Autoria própria (2020)

O motivo da escolha das lâminas das Tabelas 8 e 9, foi analisar a resposta do sistema para diferentes tonalidades, na rotina médica o número de glomérulos é importante pois é através do número de glomérulos que a lâmina é ou não analisada. As métricas foram calculadas considerando os itens de interesse (i.e., Glomérulos) e foram calculadas para cada lâmina separadamente. A figura 34 (a) mostra fragmentos das 4 lâminas utilizadas apresentadas na tabela 8. A figura 34 (b) mostra fragmentos das 3 lâminas usadas apresentadas na Tabela 8.

Figura 34 – Fragmentos das lâminas utilizadas para a avaliação de desempenho da rede, (a) mostra fragmentos das 4 lâminas utilizadas apresentadas na Tabela, (b) mostra fragmentos das 3 lâminas usadas apresentadas na Tabela 10.



Fonte: Autoria própria (2020)

4 RESULTADOS E DISCUSSÃO

Este capítulo apresenta e discute os resultados obtidos através das técnicas e ferramentas utilizadas para identificar glomérulos em imagens de lâminas inteiras de biópsia renal. Para este trabalho busca-se demonstrar quantitativamente a viabilidade de utilização de uma rede de aprendizado profundo para analisar e identificar objetos de interesse (i.e., glomérulo) em lâminas digitais inteiras de biópsia renal, utilizando como comparação resultados de métricas encontrados em literatura correlatas, e.g., (GALLEGO et al., 2018) e (LO et al., 2018).

No artigo *Glomerulus Classification and Detection Based on Convolutional Neural Networks* são utilizados dois métodos. O método I que estuda a aplicação de uma rede CNN (Rede Neural Convolucional) para classificação dos glomérulos, sendo capaz de distinguir as estruturas em glomérulo e não glomérulo. O método II estuda a função para detecção dos glomérulos nas imagens de slides inteiros e é baseado no uso de classificação de redes neurais convolucionais, para detectar as regiões dos glomérulos presentes nas imagens de lâmina inteira (GALLEGO et al., 2018). Para comparação com nossa pesquisa, vamos utilizar os valores referentes ao método II apresentado na tabela 10.

Tabela 10 – Resultados obtidos por (GALLEGO et al., 2018).

Métricas	Resultado (%)
Precisão	27,8
Sensibilidade	32,3
F1s	38,4

Fonte: GALLEGO et al. (2018)

No artigo *Glomerulus Detection on Light Microscopic Images of Renal Pathology with the Faster R-CNN* é utilizada a detecção de glomérulos de imagens inteiras de microscopia de luz total, com imagens coradas com PAS (i.e., Periodic acid-reactive Schiff) ou (H&E) (i.e., hematoxilina e eosina) obtidas com o aumento óptico de 100X, 200X e 400X (LO et al., 2018). Este artigo utiliza o faster R-CNN para a tarefa de detecção. Para a comparação com nossa pesquisa, vamos utilizar como base os valores referentes a tabela 11 extraída de (LO et al., 2018).

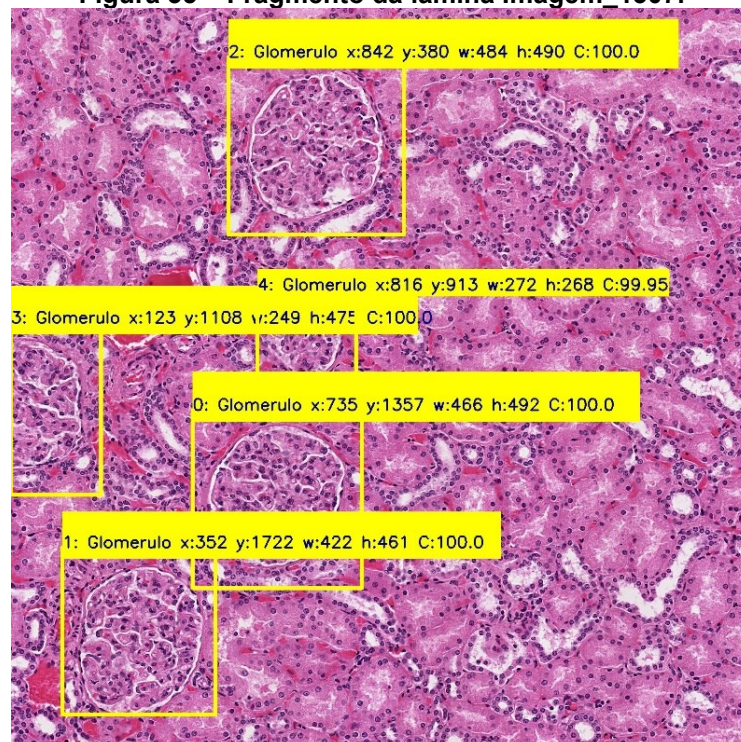
Tabela 11 – Resultados obtidos por (LO et al., 2018).

Métricas	Resultado (%)
Precisão	86,50
Sensibilidade	91,54
F1s	88,90

Fonte: LO et al. (2018)

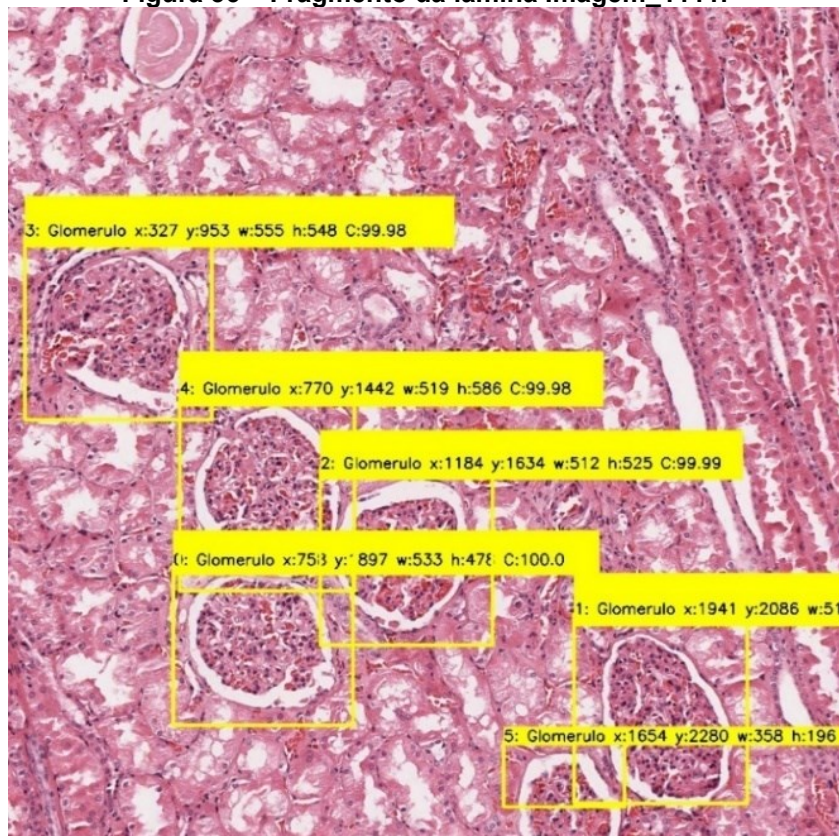
Com a utilização do framework YOLOv3 e com base nos 6 experimentos efetuados, foi possível verificar resultados superiores aos encontrados por (GALLEGO et al., 2018) e (LO et al., 2018). É importante considerar que (GALLEGO et al., 2018) realiza o processamento de classificação e detecção isoladamente e utiliza dois processamentos em paralelo, um para detecção e outro para classificação (GALLEGO et al., 2018). O mesmo caso ocorre em (LO et al., 2018) mas sem a classificação. Essa é uma das vantagens do YOLOv3, o processamento é único. Na avaliação de desempenho, é efetuada a detecção e a classificação em um único processamento. A figura 35 demonstra o fragmento de uma imagem (i.e., Imagem_1307 listada na tabela 9) do repositório BioAtlas e a figura 36 demonstra o fragmento de uma imagem (i.e., Imagem_1111 listada na tabela 10) do repositório da Universidade de Leeds, após serem processadas pela arquitetura YOLOv3 treinada na fase anterior. Em amarelo são apresentados os retângulos que marcam nas imagens as classificações dos objetos detectados apresentando a coordenada do pixel central do retângulo marcado, os valores de largura (W) e altura (H) do retângulo e a probabilidade de que o elemento marcado seja um glomérulo.

Figura 35 – Fragmento da lâmina Imagem_1307.



Fonte: Autoria própria (2020)

Figura 36 – Fragmento da lâmina Imagem_1111.



Fonte: Autoria própria (2020)

Para exemplificar o processo de análise de desempenho, a tabela 12 apresenta os valores obtidos para o Experimento E para cada lâmina (i.e., 1111, 66836 e 43267). Os índices A e B representam a metade superior e inferior da imagem de lâmina inteira, respectivamente. A tabela 13, apresenta para o mesmo experimento o cálculo das métricas.

Tabela 12 – Apresentação dos valores obtidos para o experimento E para cada lâmina (i.e., 1111, 66836 e 43267) com seus respectivos parâmetros (i.e., Tot. de imagens, Tot. Glomérulos, FP, FN, TP e Tn)

	Rep.	Img. (Qtd)	Glom. (Qtd)	(FP)	(FN)	(TP)	(TN)
Lâminas
1111_A	Leeds	295	30	2	2	27	2626
1111_B	Leeds	297	260	22	9	248	1819
Total A + B		592	290	24	11	175	4445
33863_A	Leeds	507	273	11	13	260	1917
33863_B	Leeds	507	134	15	8	126	1380
Total A + B		1014	407	26	21	386	3297
43267_A	Leeds	384	91	10	7	84	1058
43267_B	Leeds	384	151	5	5	146	1001
Total A + B		768	242	15	12	230	2059
Total	1	2374	939	65	44	891	9801

Fonte: Autoria própria (2020)

Tabela 13 – Apresentação das métricas calculadas para o experimento E.

	1111(A + B) (%)	33836 (A + B) (%)	43267 (A + B) (%)	1111 + 33836 + 43267 (%)
Métricas
Precisão	91,97	93,69	93,88	93,20
Sensibilidade	96,15	94,84	95,04	95,29
Especificidade	99,46	99,22	99,28	99,34
Eficiência	145,89	144,45	144,68	144,96
VPN	99,75	99,37	99,42	99,55
F1-Score	94,02	94,26	94,46	94,24

Fonte: Autoria própria (2020)

A tabela 14 apresenta os resultados obtidos para as métricas apresentadas em materiais e métodos, a partir da avaliação de desempenho após os treinamentos dos experimentos de A até F. Como exemplo, as Tabelas 12 e 13, apresentam os valores que foram obtidos para o Experimento E e que é apresentado de forma consolidada na coluna E da Tabela 14. Adicionalmente, para facilitar a comparação, as colunas à direita apresentam os valores obtidos por (GALLEGO et al., 2018) e (LO et al., 2018) anteriormente apresentadas nas tabelas 10 e 11.

Tabela 14 – Resultados obtidos nos experimentos de A até F efetuados com o framework YOLO utilizados todas as imagens das Tabelas 9 e 10, as duas últimas colunas são utilizadas para mostrar as informações dos artigos utilizados na comparação.

	A (%)	B (%)	C (%)	D (%)	E (%)	F (%)	GALEGO et al. (2018) (%)	LO et al. (2018) (%)
Métricas
Precisão	97,73	98,09	97,73	96,85	93,20	97,31	88,10	86,50
Sensibilidade	83,23	77,42	83,05	85,44	95,29	96,17	100,00	91,54
Acurácia	97,47	96,54	97,00	97,36	98,99	99,36
Especificidade	99,69	99,75	99,64	99,50	99,34	99,73
VPN	97,43	96,34	99,64	99,50	99,55	99,61
F1-Score	89,89	86,54	89,79	90,79	94,24	96,73	93,79	88,90

Fonte: Autoria própria (2020)

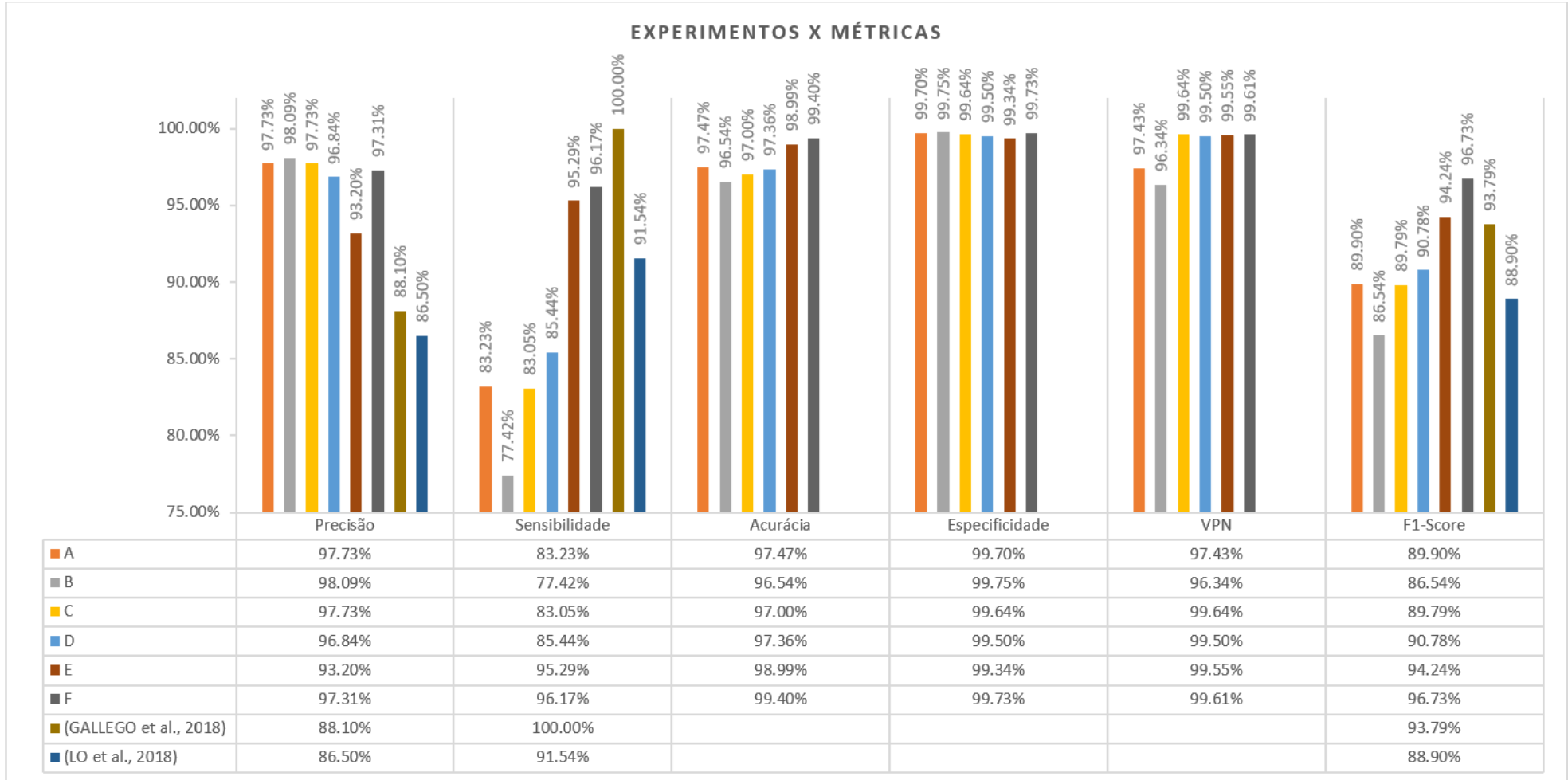
A gráfico 1 apresenta a mesma informação da tabela 14 em formato gráfico. No gráfico 1 são mostrados os valores de todas as métricas encontrados e utilizados nesta pesquisa: precisão, sensibilidade, acurácia, especificidade, VPN e F1s. Com base nestas avaliações objetivas, é possível se efetuar uma comparação entre os resultados obtidos e aqueles encontrados na literatura assim como avaliar se os mesmos apresentam potencial para a aplicação.

Os experimentos A, B, C e D tiveram o seu treinamento utilizando as imagens das tabelas 3, 4, 5 e 6, respectivamente. Nestes treinamentos, o critério de parada foi o número de iterações (i.e., aproximadamente 50.000 iterações). Nestes

experimentos, para a análise de desempenho foram utilizadas imagens de lâminas do repositório Bio-Atlas (i.e., Tabela 8). Dos resultados apresentados na tabela 14 e gráfico 1, pode-se observar que a precisão variou entre 96,84% e 98,09%, sendo o experimento B o melhor dos resultados. Para estes mesmos experimentos a acurácia variou entre 96,54% e 97,47%. De maneira similar, a especificidade variou entre 99,50% e 99,75%. Aparentemente, o uso de imagens de vários repositórios permitiu resultados robustos para a rede treinada. Ou seja, por exemplo a precisão de todos estes 4 experimentos esteve acima daquela apresentada nos trabalhos de (GALLEGO et al., 2018) e (LO et al., 2018). Por outro lado, a sensibilidade desses experimentos que variou entre 77,42% e 85,44% foi menor do que aquela apresentada por (GALLEGO et al., 2018) e (LO et al., 2018). Como a métrica F1s, combina os valores de precisão e sensibilidade, e nosso trabalho apresenta elevada sensibilidade para todos os experimentos, os valores encontrados nesta pesquisa (i.e., entre 86,54 e 90,78%) aproximam-se muito daqueles apresentados por (GALLEGO et al., 2018) e (LO et al., 2018) que são de 93,79% e 88,90%, respectivamente. Ainda em relação aos 4 primeiros experimentos, pode-se observar que o experimento D (i.e., Tabela 6) utiliza as mesmas imagens de treinamento do experimento C, mas com as imagens duplicadas. Isto permite que imagens com rotações e escalonamentos aleatórios adicionais sejam criados para o treinamento. Pode-se notar que praticamente todas as métricas do experimento D foram melhoradas, comparadas ao experimento C.

Os experimentos E e F utilizaram as imagens da tabela 3 e da tabela 7, respectivamente. Da mesma forma que os experimentos A a D, o critério de parada foi número de iterações, mas diferentemente dos experimentos A a D, utilizou o número maior de iterações (i.e., aproximadamente 60.000 iterações). Ressalta-se que o treinamento do experimento F (i.e., tabela 7) é realizado com as mesmas imagens do experimento E, mas com as imagens duplicadas permitindo um acréscimo de desempenho para todos os parâmetros com se observa na tabela 13 e no gráfico 1.

Gráfico 1 – Demonstração dos resultados dos experimentos e métricas obtidos através da tabela 13.



Fonte: Autoria própria (2020)

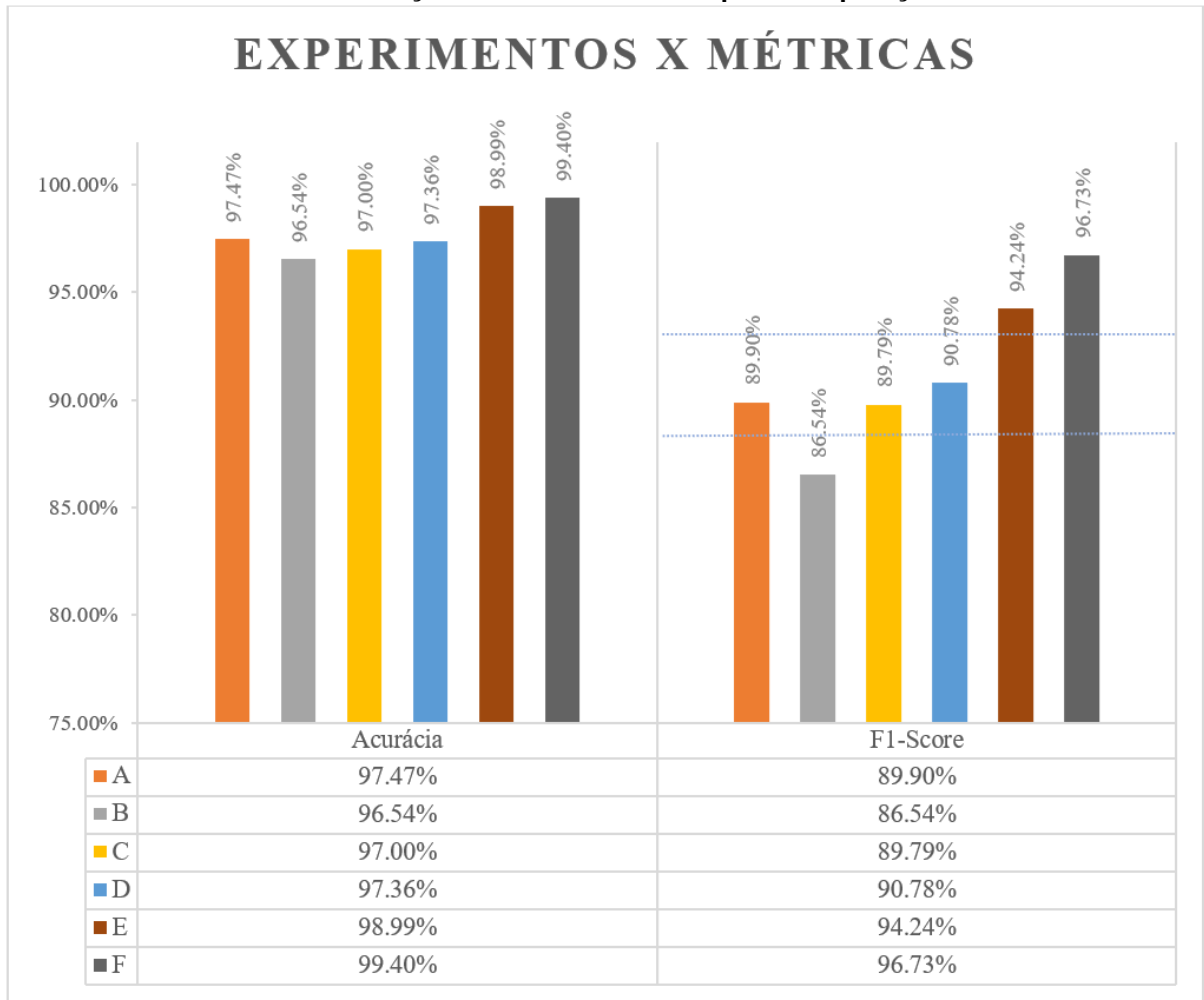
O gráfico 1 apresenta os resultados referente a acurácia e F1s dos experimentos de A até F. A acurácia representa o quão próximo da realidade estão os resultados obtidos e o F1s representa uma média entre os parâmetros de precisão e sensibilidade, onde a precisão é referente a possibilidade de uma marcação estar correta (e.g., glomérulo) e a sensibilidade é a probabilidade de uma estrutura ser encontrada (e.g., glomérulo). As tabelas 11 e 12 apresentam métricas de trabalhos recentes obtidos (GALLEGO et al., 2018) e (LO et al., 2018). Para permitir comparação com tais resultados, adicionaram-se duas linhas pontilhadas no gráfico 2 para o parâmetro de F1s. A linha superior com valor 93,79% (GALLEGO et al., 2018) e a linha inferior com o valor 88,90% (LO et al., 2018). Para os 4 primeiros experimentos, comparado com (LO et al., 2018), nossos resultados são melhores para os experimentos A, C e D. Por outro lado, os experimentos A a D apresentam pior F1s quando comparado com (GALLEGO et al., 2018). Na busca de melhores resultados foram implementados os experimentos E e F, conforme explicado anteriormente.

Para a escolha destes treinamentos, utilizamos o critério de utilizar o treinamento anterior com o melhor resultado para as métricas (i.e., acurácia e F1s), sem a duplicação das imagens no dataset, o experimento escolhido foi o A. Efetuamos um novo treinamento com este dataset até a sua finalização de forma automática pela rede YOLOv3 e chamamos de experimento E, duplicamos as imagens do dataset do treinamento E, treinamos e chamamos de experimento F.

Pode-se observar, a partir da análise do gráfico 2, que a acurácia dos experimentos E e F é 1,52% e 1,93%, respectivamente, maior que o melhor resultado apresentado entre os 4 primeiros experimentos, chegando a uma acurácia de 99,40% para o experimento F. A métrica de precisão com o valor de 97,31% apresentou um aumento de 9,20% comparado com (GALLEGO et al., 2018) e 10,81% comparado com (LO et al., 2018). Para a métrica de sensibilidade o melhor experimento (i.e., F) apresenta o valor de 96,17%. Em comparação com a literatura, temos uma redução de 3,83% e um aumento de 4,63% quando comparados com (GALLEGO et al., 2018) e (LO et al., 2018), respectivamente. Da mesma forma, a métrica F1s (i.e., média entre os parâmetros precisão e sensibilidade) foi melhorada passando de 90,78% para o melhor caso entre experimentos A a D (i.e., experimento D) para 96,73% para o experimento F. Deve-se observar que este resultado (i.e., F1s = 96,73%) é significativamente maior do que os apresentados por (GALLEGO et al., 2018) e (LO et al., 2018) com valores de 93,79% e 88,90%, respectivamente. Ou seja, um aumento em 7,83%

em relação ao estudo efetuado por (LO et al., 2018) e 2,94% em relação ao estudo efetuado por (GALLEGO et al., 2018).

Gráfico 2 – Comparação entre as métricas de acurácia e F1s de todos os experimentos analisados em relação à literatura utilizada para comparação.



Fonte: Autoria própria (2020)

5 CONCLUSÕES E TRABALHOS FUTUROS

Analisando a importância do auxílio ao diagnóstico rápido e preciso, identificou-se a necessidade e a importância de se utilizar um sistema capaz de identificar e classificar as estruturas glomerulares através de imagens de lâminas inteiras com material de biópsia renal. O sistema tem como objetivo segmentar e dar a posição de cada um dos glomérulos encontrados, nas lâminas inteiras de biópsia renal. Optou-se pela utilização de Deep Learning como método de classificação e o framework YOLOv3 que utiliza uma rede neural de 53 camadas para a realização dos experimentos.

As imagens de entrada foram padronizadas em 2048x2048 que ao serem apresentadas para a rede são redimensionadas para 512x512, parâmetro configurável, a rede processa o dataset levando aproximadamente 20 horas e gera um arquivo com os pesos gerados pelo treinamento. Este peso é apresentado para um script em Python que processa as imagens que serão apresentadas, estas imagens são processadas e devolvidas pela rede com as informações de classe e o objeto de interesse detectado.

Com base nestas imagens as informações de TP, TN, FP e FN são geradas em uma planilha do Excel. Com base nestas informações foram aplicadas as métricas de precisão, sensibilidade, acurácia, especificidade, VPN e F1s gerando os resultados apresentados nos métodos e resultados.

Com base nos resultados encontrados, podemos afirmar que o resultado do experimento F é superior aos encontrados na literatura de (GALLEGO et al., 2018) e (LO et al., 2018). Embora (GALLEGO et al., 2018) tenha na métrica sensibilidade tenha um melhor valor, chegando a 100% (i.e., A rede detecta 100% de estruturas), a métrica de precisão nos mostra que nem todas as estruturas encontradas são estruturas glomerulares (i.e., precisão = 88,10%). Isso se confirma através da métrica F1s, que mostra a média entre as métricas precisão e sensibilidade com valor de 93,79%. Este valor é inferior ao experimento e F, a diferença encontrada é de 2,94%. Assim como (GALLEGO et al., 2018), os resultados dos estudos de (LO et al., 2018) foram inferiores quando comparado ao experimento F. A diferença foi maior que as diferenças encontradas em (GALLEGO et al., 2018), a métrica sensibilidade teve o valor de 91,54% (i.e., A rede detecta 91,54% de estruturas), a métrica de precisão nos mostra que nem todas as estruturas encontradas são estruturas glomerulares (i.e., precisão

= 86,50%). Isso se confirma através da métrica F1s, que mostra a média entre as métricas precisão e sensibilidade com valor de 88,90%. Este valor (i.e., sensibilidade) é inferior ao experimento F, a diferença encontrada é de 7,83%.

O estudo efetuado, se aplicado a um sistema, pode contribuir com agilidade e segurança no auxílio à marcação e contagem das estruturas, o que facilitaria o processo de localização das estruturas de interesse (i.e., glomérulos), que posteriormente precisam ser analisadas individualmente pelos médicos patologistas. Os resultados obtidos foram satisfatórios, indicando assim que a base de dados construída também atendeu ao que foi proposto.

Como trabalho futuro, planeja-se um estudo profundo no core do framework YOLO, para que seja possível propor alterações ou melhorias no código, permitindo assim que possamos customizar e melhorar o desempenho do framework.

Buscar outras bibliotecas de software de aprendizagem profunda, como Caffe, framework de aprendizagem profunda especializado em reconhecimento de imagens, Deeplearning4j, biblioteca de aprendizagem profunda de código aberto escrita em Java/C++, TensorFlow Biblioteca de código aberto da Google para aprendizado de máquina, e outras que podemos encontrar na literatura. Considera-se também aumentar o dataset, e melhorar as anotações das imagens, com o auxílio de um especialista, para refinar as estruturas marcadas dentro dos datasets de treinamento, possibilitando assim o algoritmo possa identificar e classificar os tipos de doenças glomerulares.

Visto a importância de análise clínica, e a grande variedade de itens que podem ser analisados, sugere-se também que mais estudos sejam efetuados nesta área, com o intuito de identificar outros tecidos, células ou fluidos corporais que possam ser utilizados como experimento, possibilitando a utilização da mesma técnica de Deep Learning para identificação de características que possam ser utilizadas para auxiliar os médicos no diagnóstico dos pacientes.

REFERÊNCIAS

FACURE, Matheus. **Funções de Ativação - Entendendo a Importância Da Ativação Correta Nas Redes Neurais**. São Paulo, 2017. Disponível em: <https://matheusfacure.github.io/2017/07/12/activ-func/>. Acesso em: 13 Set. 2020.

BUENO, Gloria; FERNÁNDEZ-CARROBLES, María M. Fernández; DENIZ, Oscar; GARCÍA-ROJO, Marcial. **New Trends of Emerging Technologies in Digital Pathology**. Pathobiology, 2016, v. 83, No. 2-3, p 61–69. DOI: <https://doi.org/10.1159/000443482>. Disponível em: <https://www.karger.com/Article/FullText/443482>.

CHICCO, Davide; JURMAN, Giuseppe. **The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation**. BMC Part of Springer Nature 2020, p: 1–13. DOI: <https://doi.org/10.1186/s12864-019-6413-7>. Disponível em: <https://bmcbgenomics.biomedcentral.com/track/pdf/10.1186/s12864-019-6413-7.pdf>.

PATHOLOGY, Dept, Virtual. **Virtual Pathology at the University of Leeds**. Division of Pathology & Data Analytics, Leeds Institute of Medical Research, University of Leeds. Disponível em: <https://www.virtualpathology.leeds.ac.uk/>. Acessado em: 10 julh. 2020.

DOUGHERTY, Geoff. **Digital Image Processing for Medical Applications**. California. Cambridge University Press. 2009 ISBN: 9780521860857.

EUSEBI, Paolo. **Diagnostic Accuracy Measures**. Cerebrovascular Diseases, 2013, v. 36, No. 4, p. 267–72 DOI: <https://doi.org/10.1159/000353863> Disponível em: <https://www.karger.com/Article/FullText/353863>.

TEMERINAC-OTT, M.; FORESTIER, G.; SCHMITZ, J.; HERMSEN, M.; BRÄSEN, J. H.; FEUERHAKE, F.; WEMMERT, C. **Detection of Glomeruli in Renal Pathology by Mutual Comparison of Multiple Staining Modalities**. Proceedings of the 10th International Symposium on Image and Signal Processing and Analysis, 2017, p. 19–24, DOI: <https://doi.org/10.1109/ISPA.2017.8073562>.

FOUNDATION, Jake Gittlen Cancer Research; PA Tobacco Settlement Fund. **Bio-Atlas at the Jake Gittlen Laboratories for Cancer Research**. NIH grant 5R24 RR01744, Disponível em: <http://bio-atlas.psu.edu/>. Acessado em: 10 julh. 2020.

GALLEGO, J.; PEDRAZA, A.; LOPEZ, S.; STEINER, G.; GONZALEZ, L.; LAURINAVICIUS, A.; BUENO, G. **Glomerulus Classification and Detection Based on Convolutional Neural Networks**. Journal of Imaging. 2018, v. 4 , p. 20. Disponível em: <http://www.mdpi.com/2313-433X/4/1/20>. DOI: <https://doi.org/10.3390/jimaging4010020>.

GONZALEZ, Rafael C.; WOODS, Richard C. **Processamento Digital de Imagens 3**. Ed. Softcover. Pearson, 2009. ISBN: 978-8-5760-5401-6.

GUYTON, Arthur C.; HALL John E. **Tratado de Fisiologia Médica**. 12. ed. Rio de

Janeiro: Elsevier. 2011. ISBN: 978-1-4160-4574-8 .

HINTON, Geoffrey E. **Learning Multiple Layers of Representation**. Trends in Cognitive Sciences 2007. v. 11, p. 428–34.

JANOWCZYK, Andrew; MADABHUSHI, Anant. **Deep Learning for Digital Pathology Image Analysis : A Comprehensive Tutorial with Selected Use Cases**. Journal of Pathology Informatics. 2016 v. 7, p. 29, DOI: <https://doi.org/10.4103/2153-3539.186902>.

JUNQUEIRA, Luiz Carlos Uchoa; CARNEIRO, José. **Histologia Básica Texto e Atlas 12º Edição**. 12. ed. Rio de Janeiro: Editora Guanabara Koogan LTDA. 2013.

KATO, Tsuyoshi; RELATOR, Raissa; NGOUV, Hayliang; HIROHASHI, Yoshihiro; TAKAKI, Osamu; KAKIMOTO, Tetsuhiro; OKADA, Kinya. **Segmental HOG: New Descriptor for Glomerulus Detection in Kidney Microscopy Image**. BMC Part of Springer Nature, 2015 v. 16, p. 1–16, DOI: <https://doi.org/10.1186/s12859-015-0739-1>

KAWAMURA, Takao. **Interpretação de Um Teste Sob a Visão Epidemiológica. Eficiência de Um Teste**. São Paulo, Araçatuba, 2002, Scielo – Arquivos Brasileiros de Cardiologia v. 79, no. 4, Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0066-782X2002001300015.

KREUTZER, Ralf T.; SIRRENBURG, Marie. **Understanding Artificial Intelligence. Fundamentals, Use Cases and Methods for a Corporate AI Journey**, Suíça: Springer Nature, 2020. p. 1-273. DOI: <https://doi.org/10.1007/978-3-030-25271-7>.

KUMAR, Vinay; ABBAS, Abul K.; ASTER, Jon C. **Robbins Patologia Básica. 9º. 9 Ed.** Rio de Janeiro, Elsevier, 2013, p.928, ISBN: 978-1-4377-1781-5.

LEICA. **Leica Biosystems**. Advancing Cancer Diagnostics Improving Lives. Disponível em: <https://www.leicabiosystems.com/digital-pathology/manage/aperio-imagescope/>. Acessado em: 01 Jan. 2020

LITJENS, Geert; SÁNCHEZ, Clara I.; TIMOFEEVA, Nadya; HERMSEN, Meyke; NAGTEGAAL, Iris; KOVACS, Iringo; KAA, Christina Hulsbergen-Van; BULT, Peter; VAN GINNEKEN, Bram; VAN DER LAAK, Jeroen. **Deep Learning as a Tool for Increased Accuracy and Efficiency of Histopathological Diagnosis**. Scientific Reports, v. 6, no. 26286, DOI: <https://doi.org/10.1038/srep26286>

LO, Ying Chih; JUANG, Chia Feng; CHUNG, I. Fang; GUO, Shin Ning; HUANG, Man Ling; WEN, Mei Chin; LIN, Cheng Jian; LIN, Hsueh Yi. **Glomerulus Detection on Light Microscopic Images of Renal Pathology with the Faster R-CNN**. Neural Information Processing. ICONIP 2018. Lecture Notes in Computer Science, v. 11307, p. 369-377, DOI: https://doi.org/10.1007/978-3-030-04239-4_33.

MAAS, Andrew L; HANNUN, Awni Y.; NG, Andrew Y. **Rectifier Nonlinearities Improve Neural Network Acoustic Models**. in *ICML Workshop on Deep Learning for Audio, Speech and Language Processing*, California, 2013

MORETTI, Everson; HASSUNUMA, Renato Massaharu; GARCIA, Patricia Carvalho; MESSIAS, Sandra Heliísa Nunes. **Guia Visual de Processamento Histológico**. 1 ed. São Paulo: Canal6editora 2019. Disponível em: http://www.canal6.com.br/livros_loja/Ebook_Guia_Visual_Processamento_Histologico.pdf . ISBN: 978-85-7917-557-2

NETO, Hugo Vieira; FILHO, Ogê Marques. **Processamento Digital de Imagem**. Rio de Janeiro: Brasport. 1999. ISBN 8574520098.

OBERHOLZER, Martin; ÖSTREICHER, Marc; CHRISTEN, Heinz; BRÜHLMANN, Marcel. **Methods in Quantitative Image Analysis**. Histochemistry and Cell Biology, 1996. v.105, p. 333–55 DOI: <https://doi.org/10.1007/BF01463655>.

MARÉE, Raphaël; DALLONGEVILLE, Stéphane; OLIVIO MARIN J.C.; MEAS-YEDID, Vannary. **An Approach for Detection of Glomeruli in Multisite Digital Pathology**. IEEE 13th International Symposium on Biomedical Imaging ISBI 2016, p.1033–36 DOI: 10.1109/ISBI.2016.7493442.

REDMON, Joseph; DIVVALA, Santosh; GIRSHCK, Ross; FARHADI, Ali. **You Only Look Once: Unified, Real-Time Object Detection Joseph**. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition – CVPR, 2016, p. 779-788

REDMON, Joseph; FARHADI Ali **YOLOv3: An Incremental Improvement**. Cornell University 2018, v.1, Disponível em: <http://arxiv.org/abs/1804.02767> ou <https://arxiv.org/abs/1804.02767v1>.

ROSAI, Juan. **Juan Rosai's Collection of Surgical Pathology Seminars (1945 - Present)**. Disponível em: <https://www.rosaicollection.org/aboutRosai.cfm>. Acessado em 20, julh 2020.

RUSS, John c.; NEAL, F. Brent. **The Image Processing Handbook**. 7 ed. Taylor & Francis Group NW, 2016. ISBN: 978-11-387-4749-4.

RUSSELL, Stuart; NORVIG, Peter. **Artificial Intelligence: A Modern Approach**. 4 ed. Pearson. Hoboken NJ 2020. ISBN: 978-01-346-1099-3.

SILAPARASETTY, Vinita. **Deep Learning Projects Using TensorFlow 2: Neural Network Development with Python and Keras**. India, 2020. ISBN: 978-1-4842-5802-6.

SOLOMON, Chris; BRECKON, Toby. **Fundamentos de Processamento Digital de Imagens: Uma abordagem Prática com Exemplos em Matlab**. Tradução e Revisão técnica: J. R. Souza. 1 ed. Rio de Janeiro, LTC - GRUPO GEN, 2013. ISBN: 978-85-216-2347-2

SUN, Draining. **YBAT-YOLO BBox Annotation Tool**. Disponível em: <https://github.com/drainingsun/ybat>. Acessado em 01 Mar. 2020.

TAULLI, Tom. **Introdução à Inteligência Artificial: Uma Abordagem Não Técnica**. Tradução: Luciana do Amaral Teixeira. São Paulo: Novatec, 2020. 1 ed. ISBN: 978-85-7522-819-7.

TORTORA, Gerard J.; NIELSEN, Mark T. **Princípios de Anatomia Humana**. 12 ed.

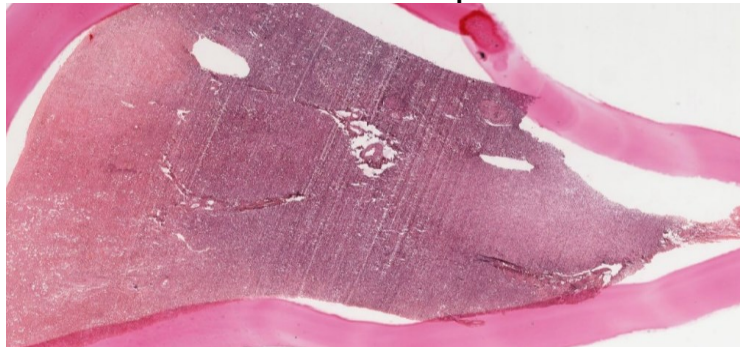
Rio de Janeiro, 2013. Editora Guanabara Koogan LTDA. ISBN: 978-85-277-2065-6.

TURING, Alan M. **Computing Machinery and Intelligence**. Disponível em: <https://archive.nytimes.com/www.nytimes.com/library/cyber/surf/1106surf-turing.html>. Acessado em 08 Jan. 2020.

APÊNDICE A - Bases de imagens Juan Rosai's

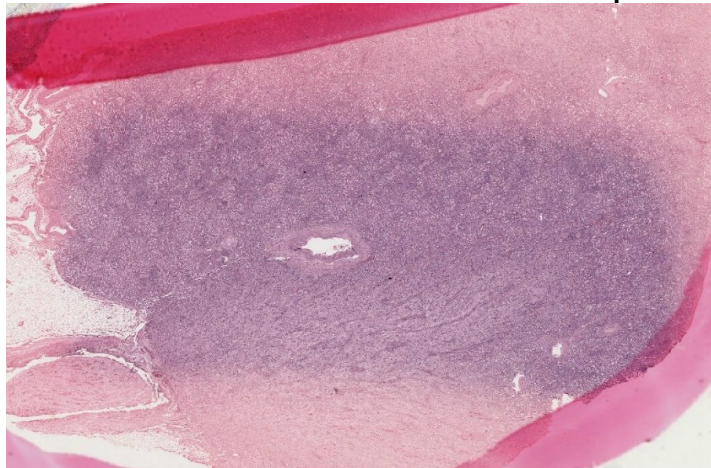
Visto a necessidade de imagens inteiras de biópsia renal para dar sequência ao estudo, buscamos na internet base de imagens que nos disponibilizassem uma grande quantidade de imagens, permitindo que o treinamento da rede fosse efetuado. As figuras 37 a 43, representam as 7 lâminas de rins humanos utilizadas no treinamento e validação da rede do framework YOLOv3, retiradas da base de dados Juan Rosai's.

Figura 37 – Imagem caso 8 utilizada para treinamento tamanho em pixel 45967x26126 arquivo de 3.36GB podendo ser aproximara em até 20x, <http://rosai.secondslide.com/sem131/sem131-case8.svs/view.apml>.



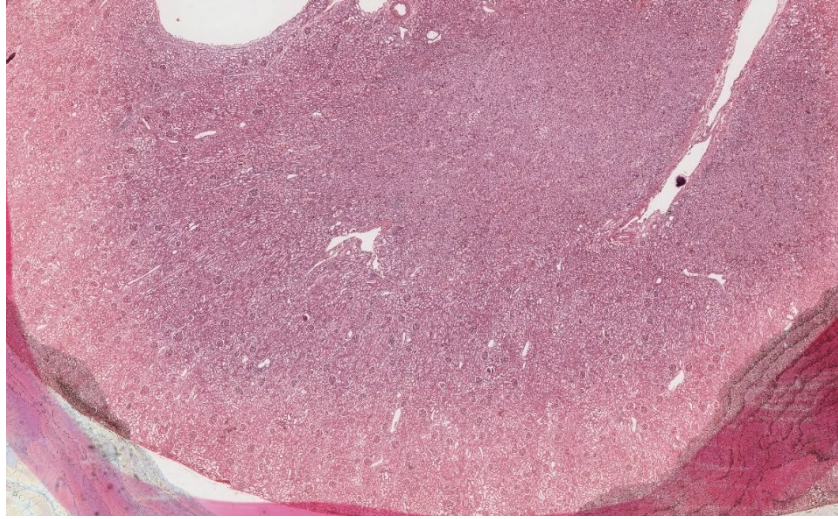
Fonte: Aatoria própria (2020)

Figura 38 – Imagem caso 10 utilizada para o treinamento da rede, tamanho em pixel 37980x27396 arquivo de 2.91GB, pode ser aproximado em até 20x, <http://rosai.secondslide.com/sem131/sem131-case10.svs/view.apml>.



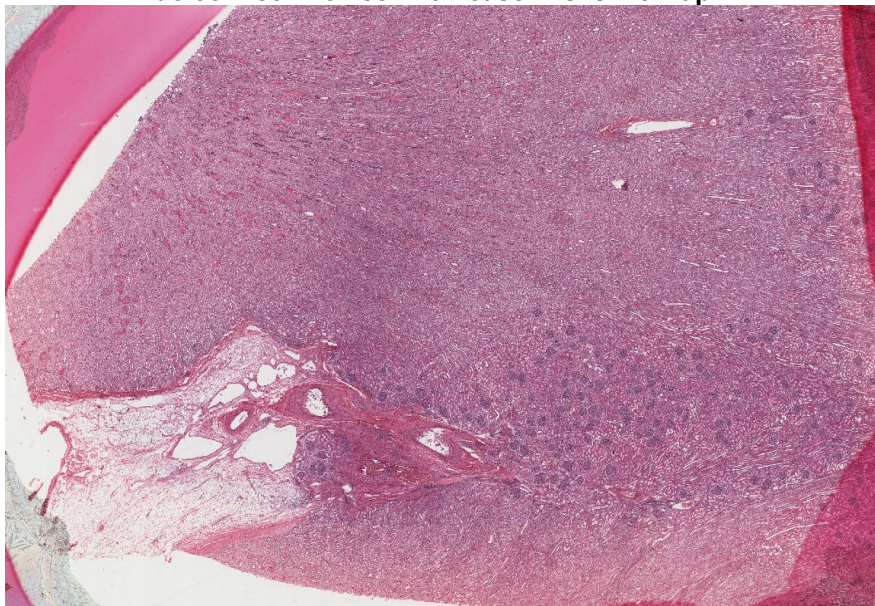
Fonte: Aatoria própria (2020)

Figura 39 – Imagem caso 11 utilizada para o treinamento da rede, tamanho em pixel 37980x24075 arquivo de 2.55GB, pode ser aproximado em até 20x, <http://rosai.secondslide.com/sem131/sem131-case11.svs/view.apml>.



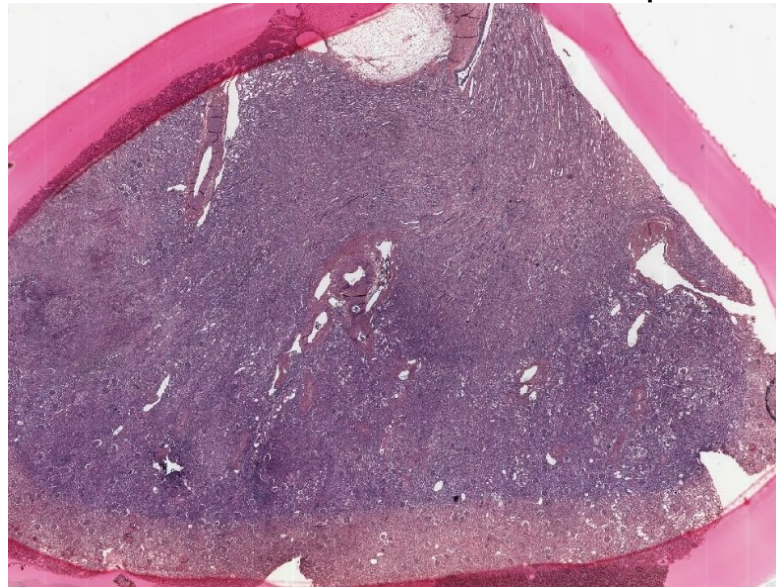
Fonte: A autoria própria (2020)

Figura 40 – Imagem caso 12 utilizada para o treinamento da rede, tamanho em pixel 35981x23562 arquivo de 2.37GB, pode ser aproximado em até 20x, <http://rosai.secondslide.com/sem131/sem131-case12.svs/view.apml>.



Fonte: A autoria própria (2020)

Figura 41 – Imagem caso 13 utilizada para o treinamento da rede, tamanho em pixel 37980x29313 arquivo de 3.11GB, pode ser aproximado em até 20x, <http://rosai.secondslide.com/sem131/sem131-case13.svs/view.apml>.



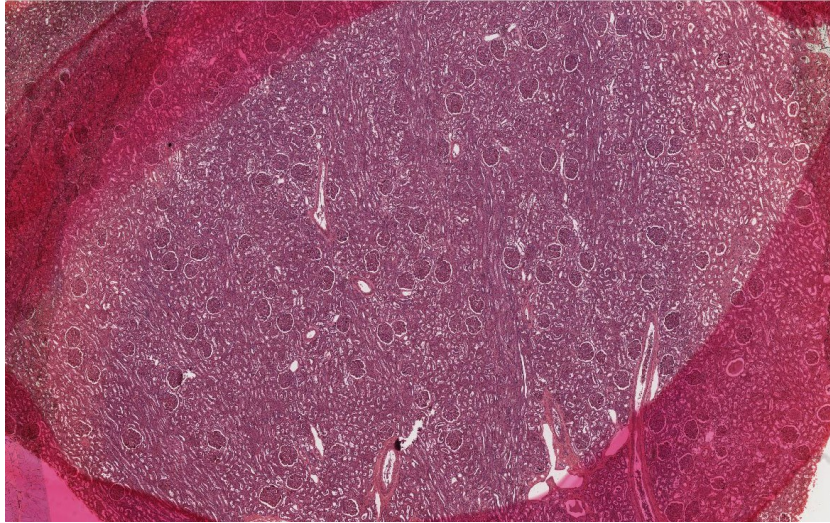
Fonte: Autoria própria (2020)

Figura 42 – Imagem caso 14 utilizada para o treinamento da rede, tamanho em pixel 33982x23944 arquivo de 2.27GB, pode ser aproximado em até 20x, <http://rosai.secondslide.com/sem131/sem131-case14.svs/view.apml>.



Fonte: Autoria própria (2020)

Figura 43 – Imagem caso 15 utilizada para o treinamento da rede, tamanho em pixel 19989x12818 arquivo de 733.05MB, pode ser aproximado em até 20x, <http://rosai.secondslide.com/sem131/sem131-case15.svs/view.apml>.

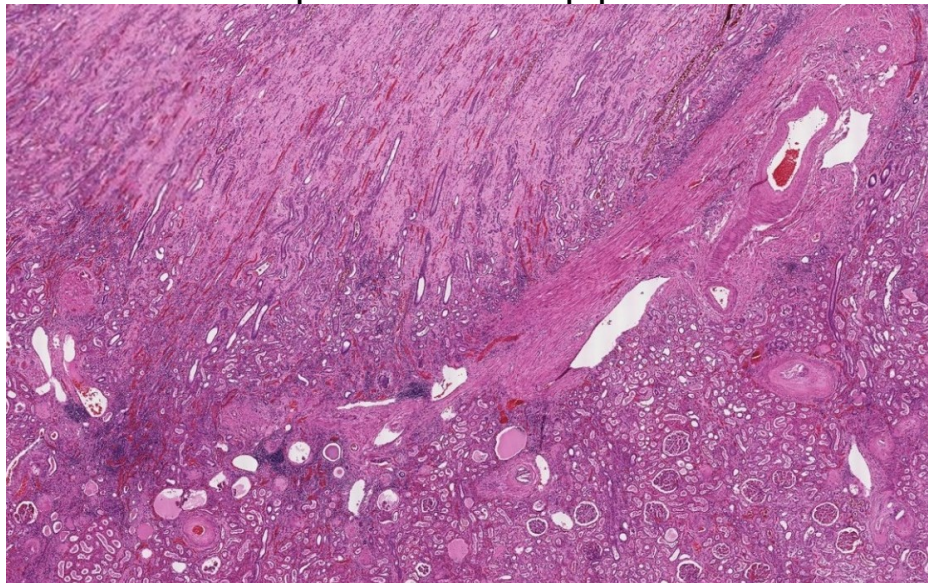


Fonte: Autoria própria (2020)

APÊNDICE B - Bases de imagens Bio-Atlas

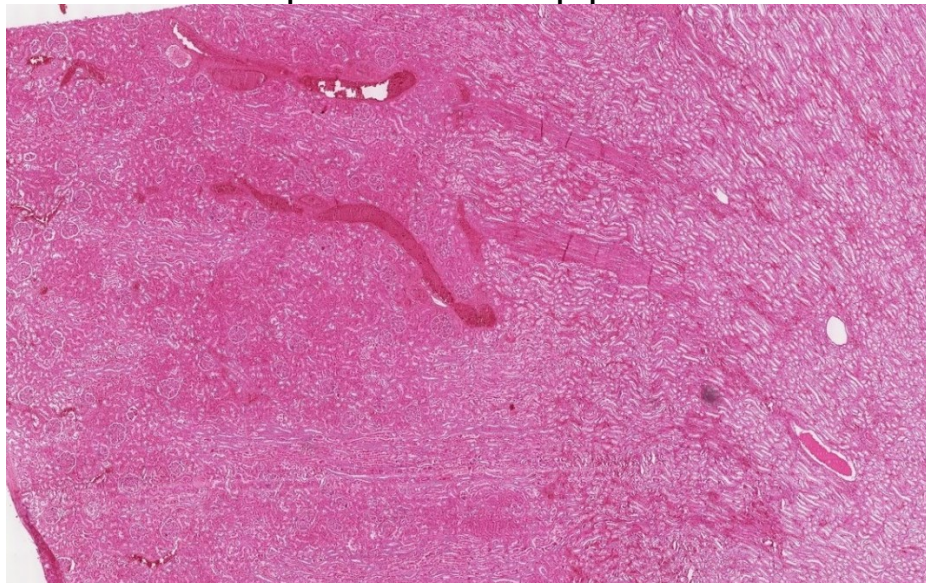
A base de imagens Bio-Atlas forneceu imagens tanto para o treinamento e validação da rede YOLOv3, quanto para a avaliação de desempenho. As figuras 44 a 48, representam as 5 lâminas de rins humanos utilizadas no treinamento e validação da rede do framework YOLOv3. Já as figuras de 49 a 53, representam as 5 lâminas inteiras utilizadas para a avaliação de desempenho da rede (i.e., processo executado após o treinamento da rede)

Figura 44– Imagem 837 imagem de biópsia renal utilizada para treinamento da rede, <http://bio-atlas.psu.edu/human/view.php?s=837>.



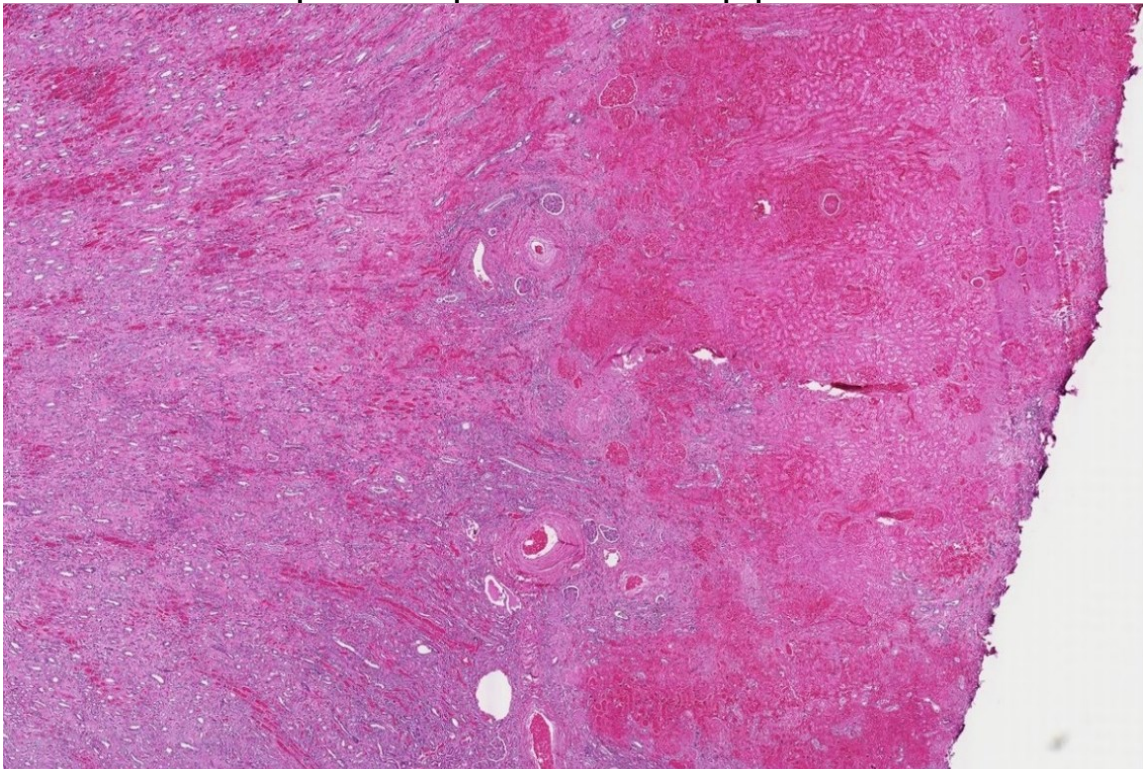
Fonte: Autoria própria (2020)

Figura 45– Imagem 1060 imagem de biópsia renal utilizada para treinamento da rede, <http://bio-atlas.psu.edu/human/view.php?s=1060>.



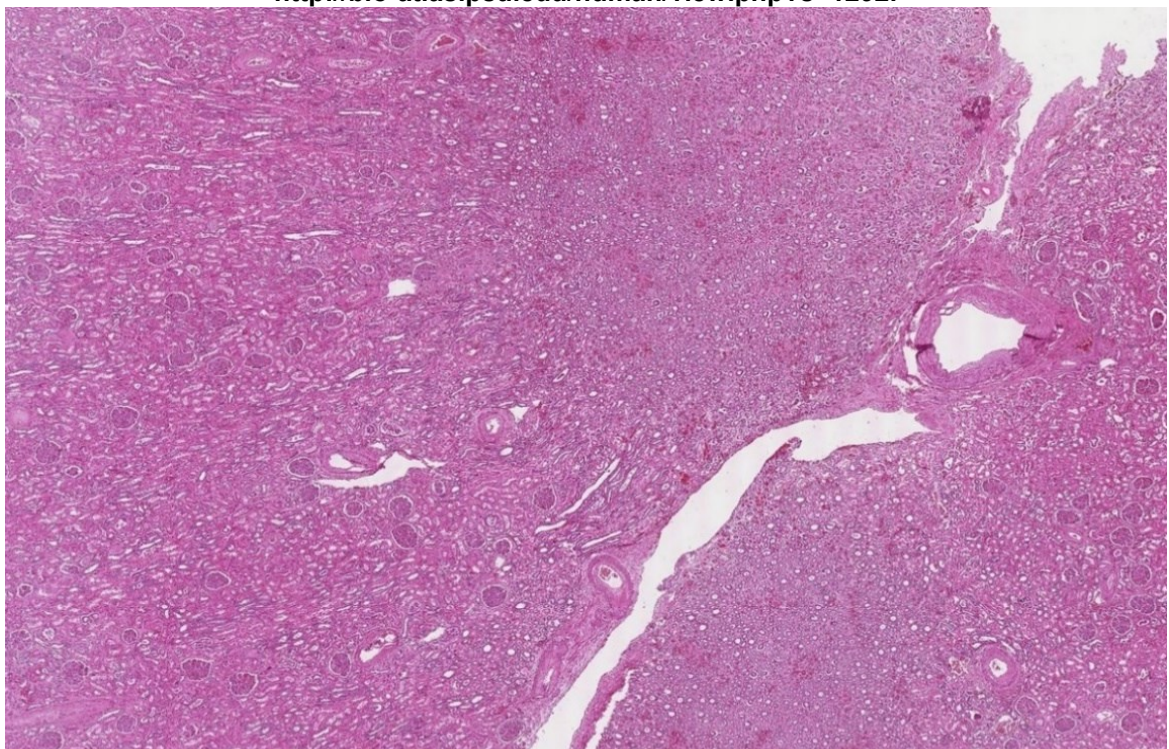
Fonte: Autoria própria (2020)

Figura 46 – Imagem 1267 imagem de biópsia renal utilizada para treinamento da rede, <http://bio-atlas.psu.edu/human/view.php?s=1267>.



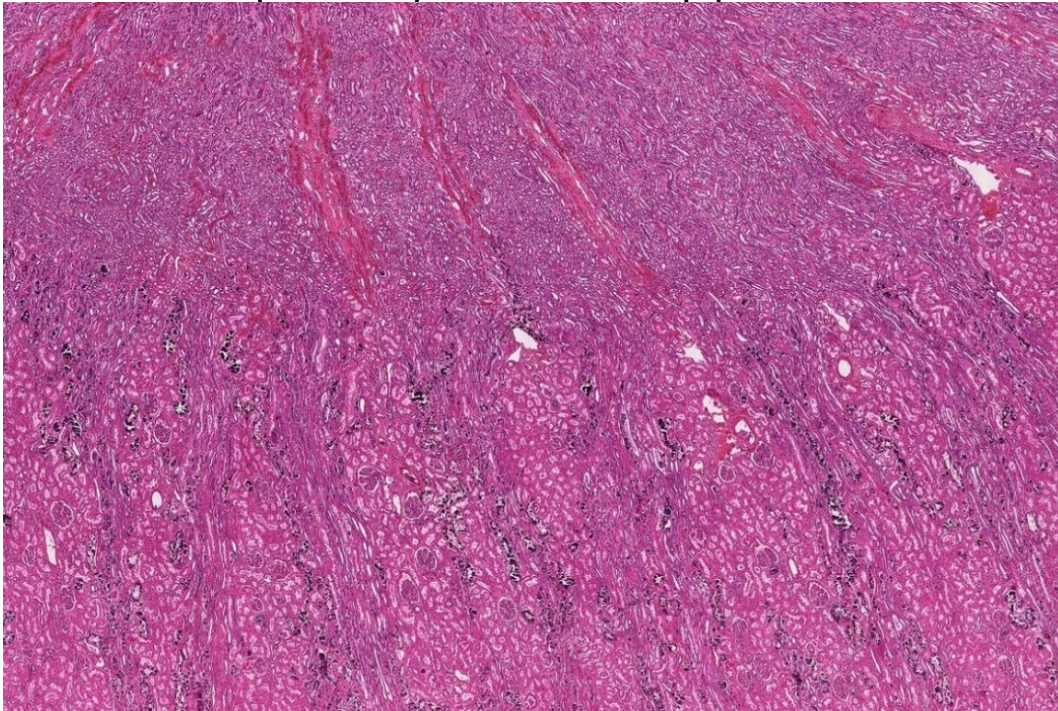
Fonte: Autoria própria (2020)

Figura 47 – Imagem 1292 imagem de biópsia renal utilizada para treinamento da rede, <http://bio-atlas.psu.edu/human/view.php?s=1292>.



Fonte: Autoria própria (2020)

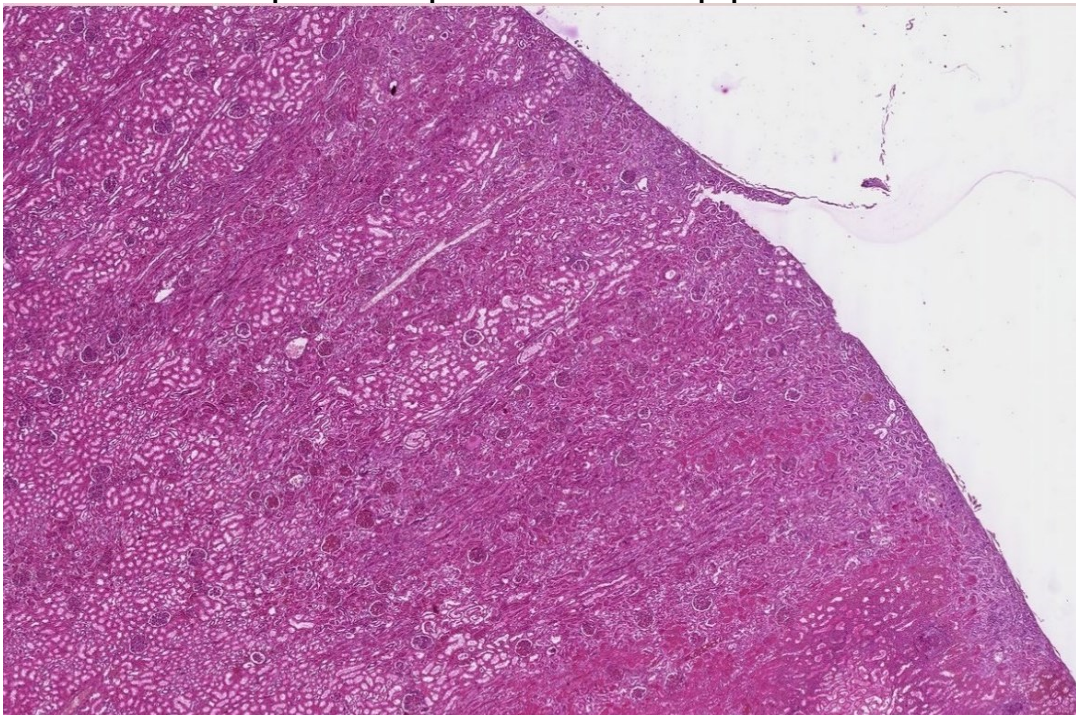
Figura 48– Imagem 1298 imagem de biópsia renal utilizada para validação do treinamento, <http://bio-atlas.psu.edu/human/view.php?s=1298>.



Fonte: A autoria própria (2020)

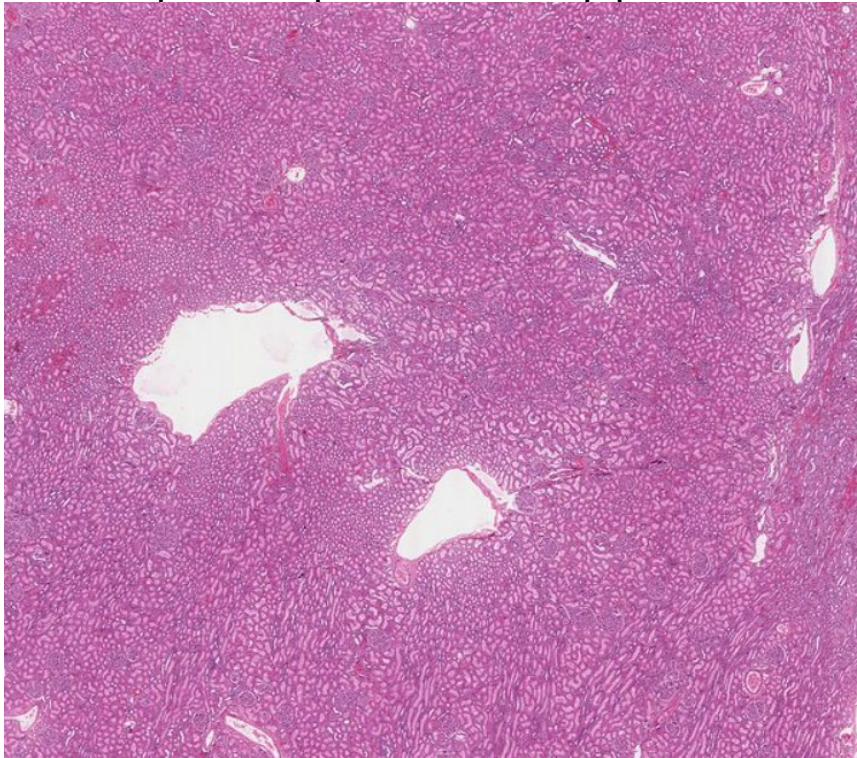
Como citado anteriormente, as Figuras 49 a 53, mostram as lâminas utilizadas para a avaliação de desempenho da rede YOLOv3.

Figura 49 – Imagem 569 imagem de biópsia renal utilizada para avaliação de desempenho, <http://bio-atlas.psu.edu/human/view.php?s=569>.



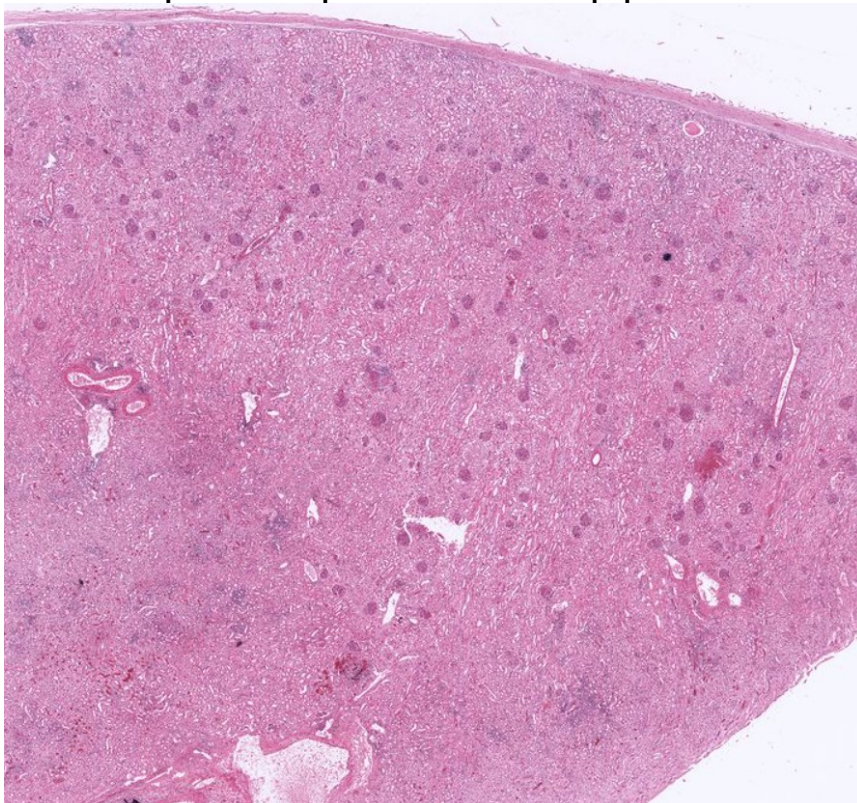
Fonte: A autoria própria (2020)

Figura 50 – Imagem 1307 imagem de biópsia renal utilizada para avaliação de desempenho, <http://bio-atlas.psu.edu/human/view.php?s=1307>.



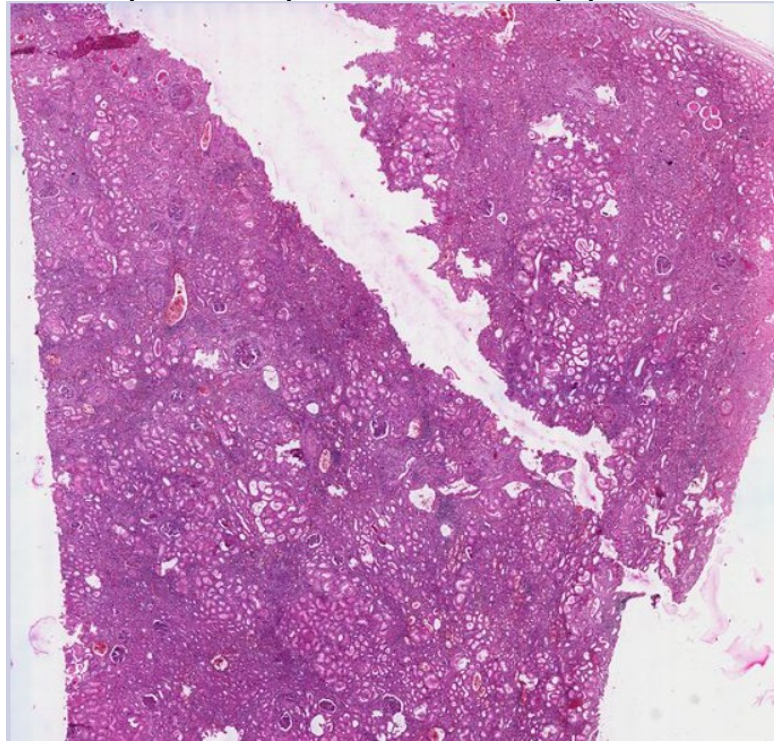
Fonte: Autoria própria (2020)

Figura 51 – Imagem 603 imagem de biópsia renal utilizada para avaliação de desempenho, <http://bio-atlas.psu.edu/human/view.php?s=603>.



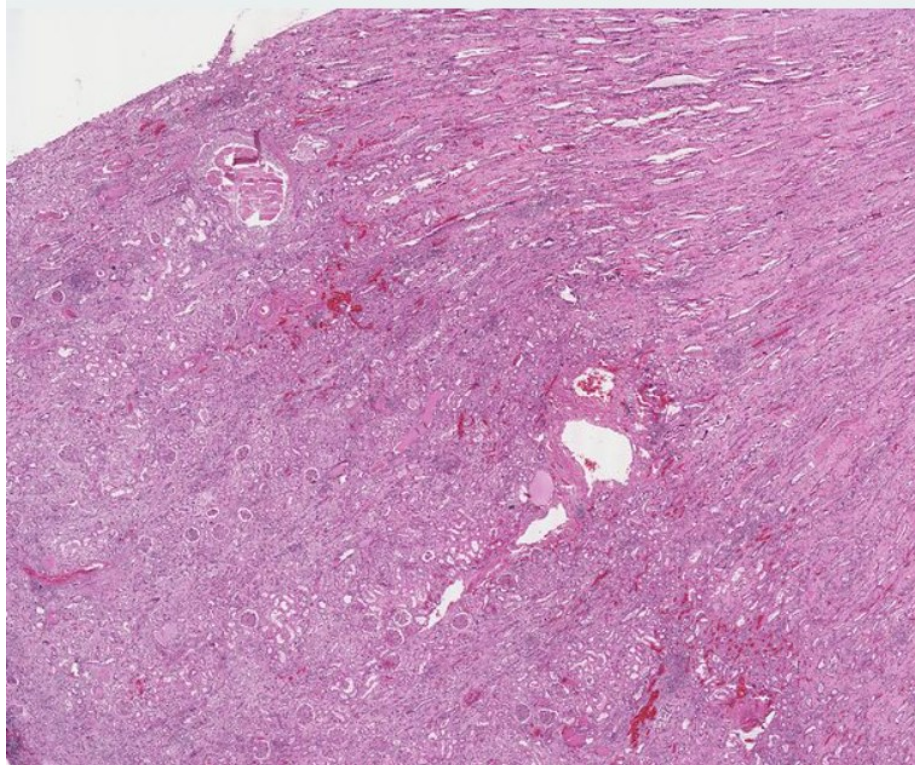
Fonte: Autoria própria (2020)

Figura 52 – Imagem 603 imagem de biópsia renal utilizada para avaliação de desempenho, <http://bio-atlas.psu.edu/human/view.php?s=603>



Fonte: Autoria própria (2020)

Figura 53 – Imagem 849 imagem de biópsia renal utilizada para avaliação de desempenho, <http://bio-atlas.psu.edu/human/view.php?s=849>.

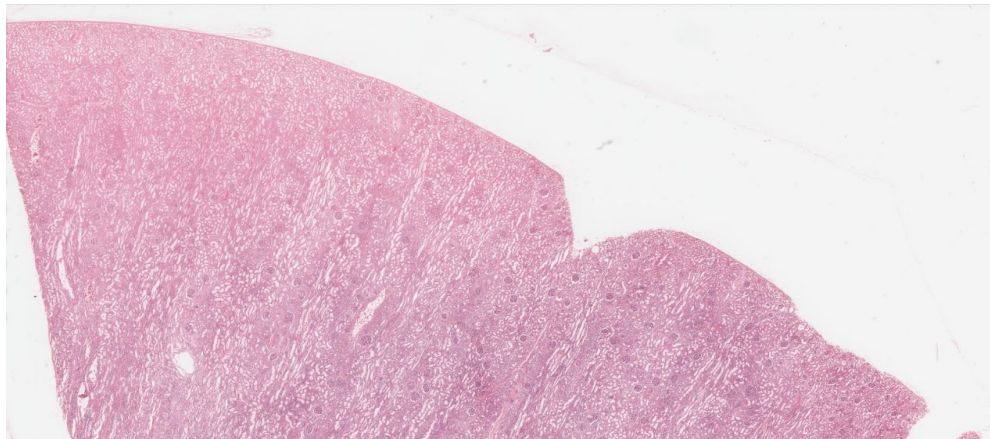


Fonte: Autoria própria (2020)

APÊNDICE C - Bases de imagens LEEDS

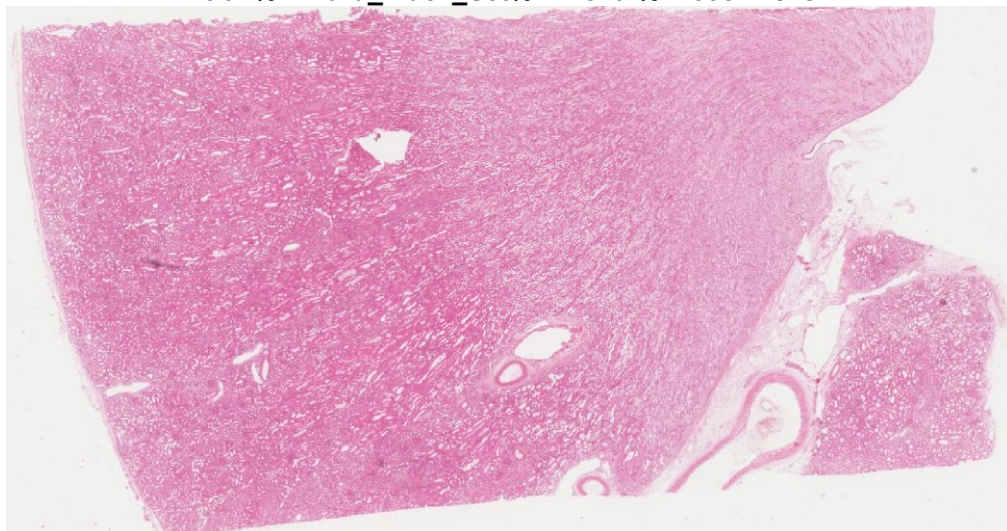
A base de imagens Leeds teve um processo similar a base Bio-atlas, essa base forneceu também, imagens tanto para o treinamento e validação da rede YOLOv3, quanto para a avaliação de desempenho. As Figuras 54 a 57, representam as 4 lâminas de rins humanos utilizadas no treinamento e validação da rede do framework YOLOv3. As Figuras de 75 a 77, representam as 3 lâminas inteiras utilizadas para a avaliação de desempenho da rede (i.e., processo executado após o treinamento da rede)

Figura 54– Apresenta a lâmina completa 1057 utilizada para o treinamento da rede, tamanho em pixel 48024x35461, https://www.virtualpathology.leeds.ac.uk/slides/library/view.php?path=%2FResearch_4%2FTeaching%2FEducation%2FUndergraduate%2F1057.svs.



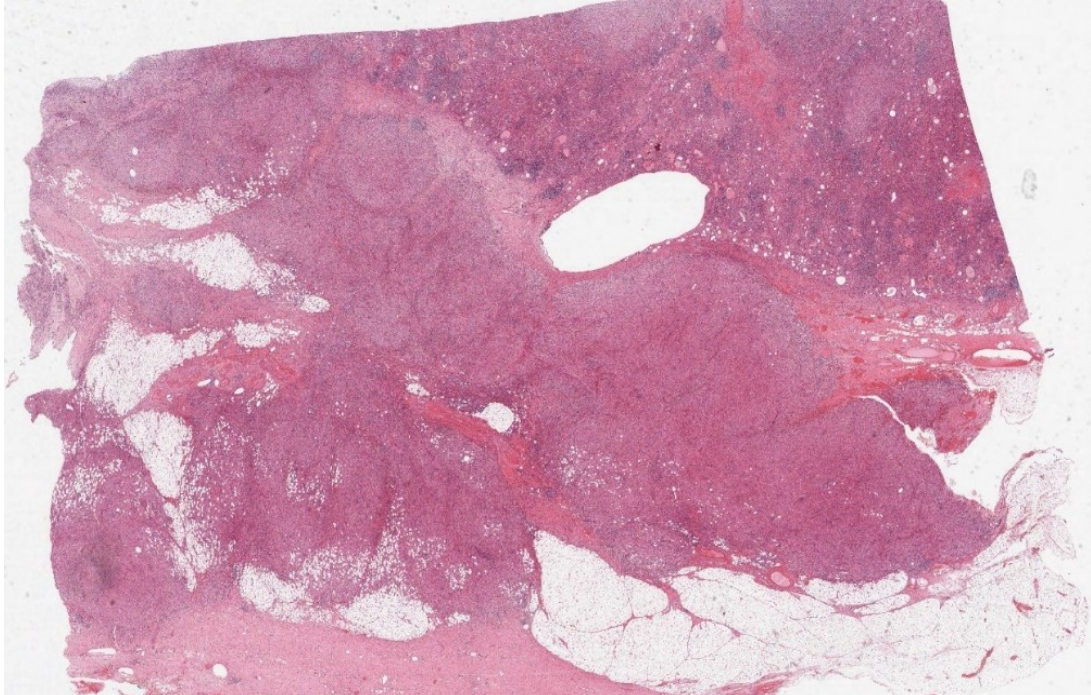
Fonte: Autoria própria (2020)

Figura 55 – Apresenta a lâmina completa 33814 utilizada para o treinamento da rede, tamanho em pixel 81600x47466, https://www.virtualpathology.leeds.ac.uk/slides/browser/view_flash.php?path=%2FResearch_4%2FSlide_Library%2FR_Bishop_Collection%2FCard_index_Set%2FRenal%2F33814.svs



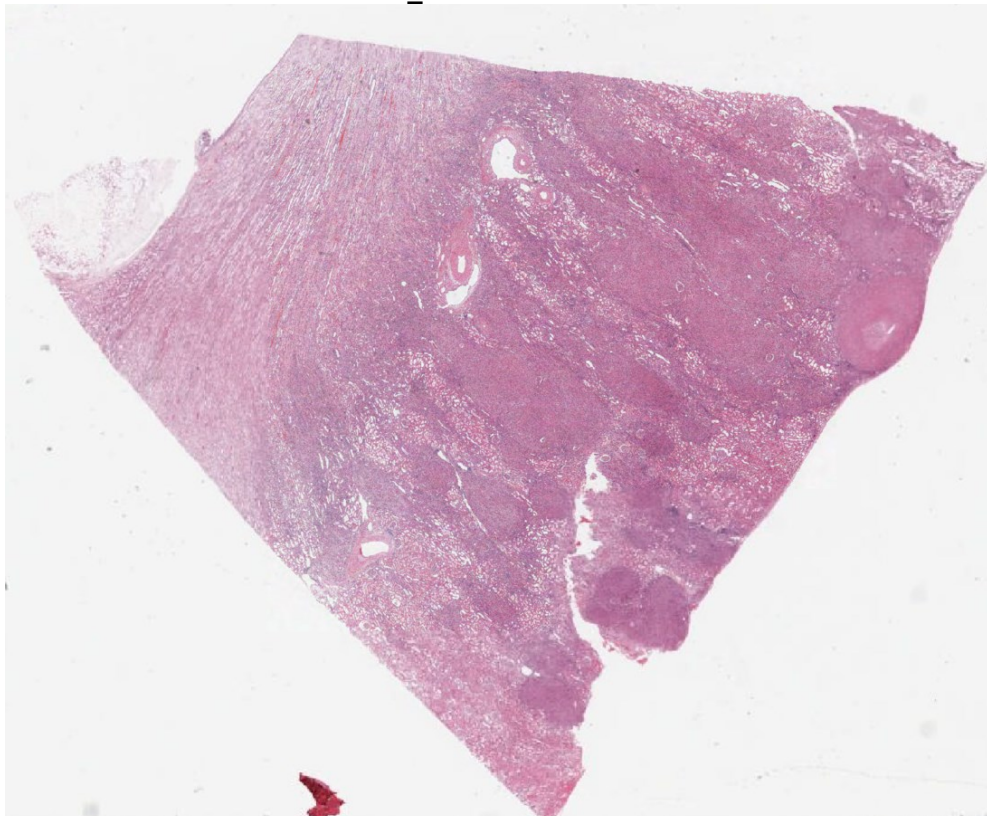
Fonte: Autoria própria (2020)

Figura 56 – Apresenta a lâmina completa 134159 utilizada para o treinamento da rede, tamanho em pixel 54263x40252, <https://tinyurl.com/yxqxq9gr>.



Fonte: Autoria própria (2020)

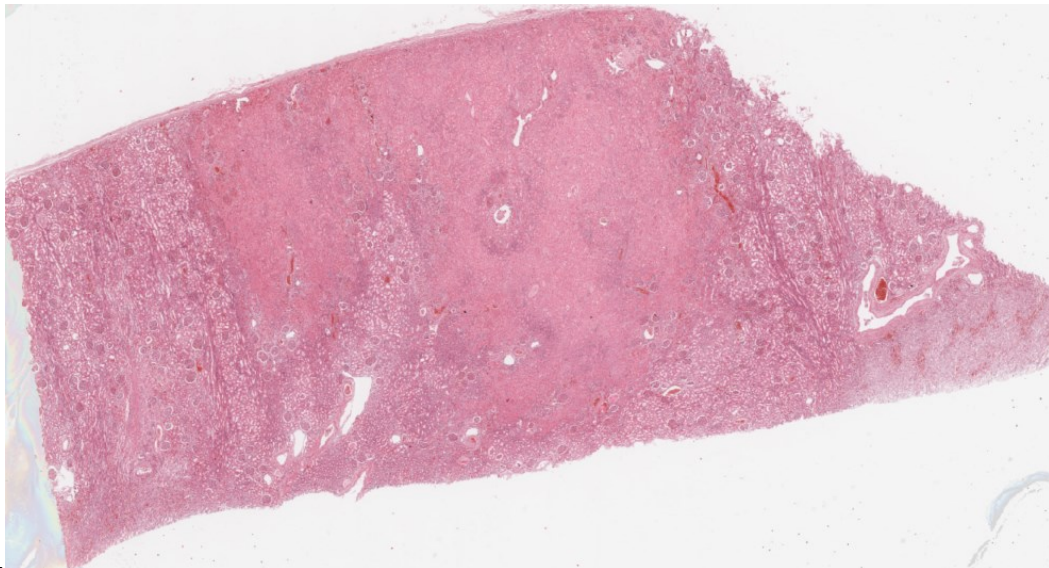
Figura 57 – Apresenta a lâmina completa 149335 utilizada para o treinamento da rede, tamanho em pixel 66639x48640, https://www.virtualpathology.leeds.ac.uk/slides/library/view.php?path=%2FResearch_4%2FTeaching%2FEQA%2FEast_Midlands%2FCirculation_G%2F149335.svs



Fonte: Autoria própria (2020)

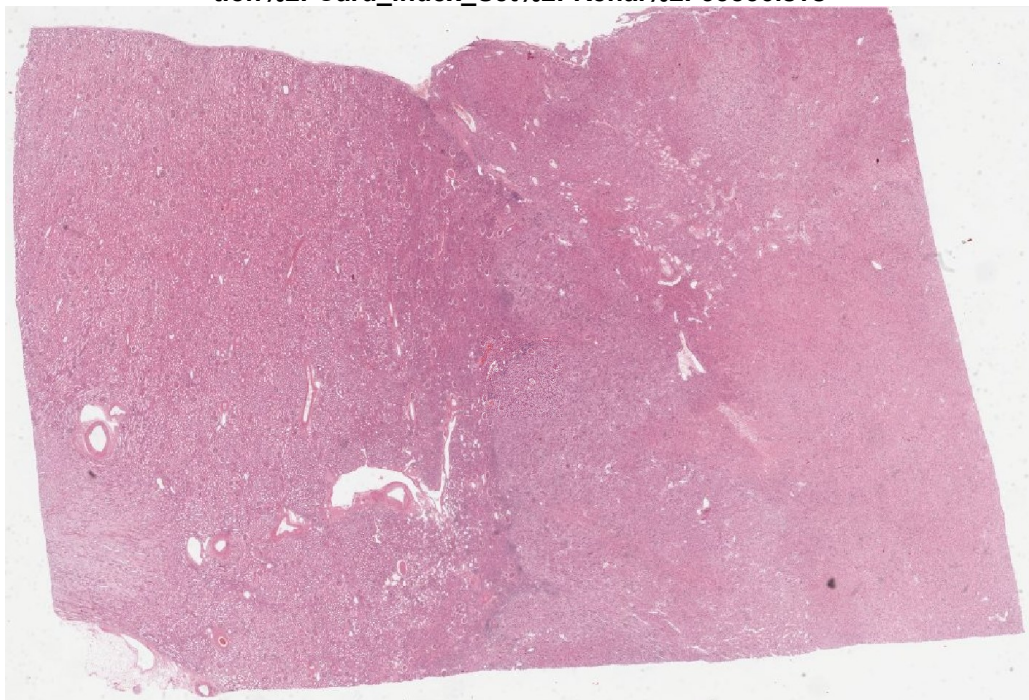
Como citado anteriormente, as figuras 58 a 60, mostram as lâminas utilizadas para a avaliação de desempenho da rede YOLOv3.

Figura 58 – Apresenta a lâmina completa 1111 utilizada para a avaliação de desempenho da rede, tamanho em pixel 82041x45123, https://www.virtualpathology.leeds.ac.uk/slides/library/view.php?path=%2FResearch_4%2FTeaching%2FEducation%2FUndergraduate%2F1111.svs



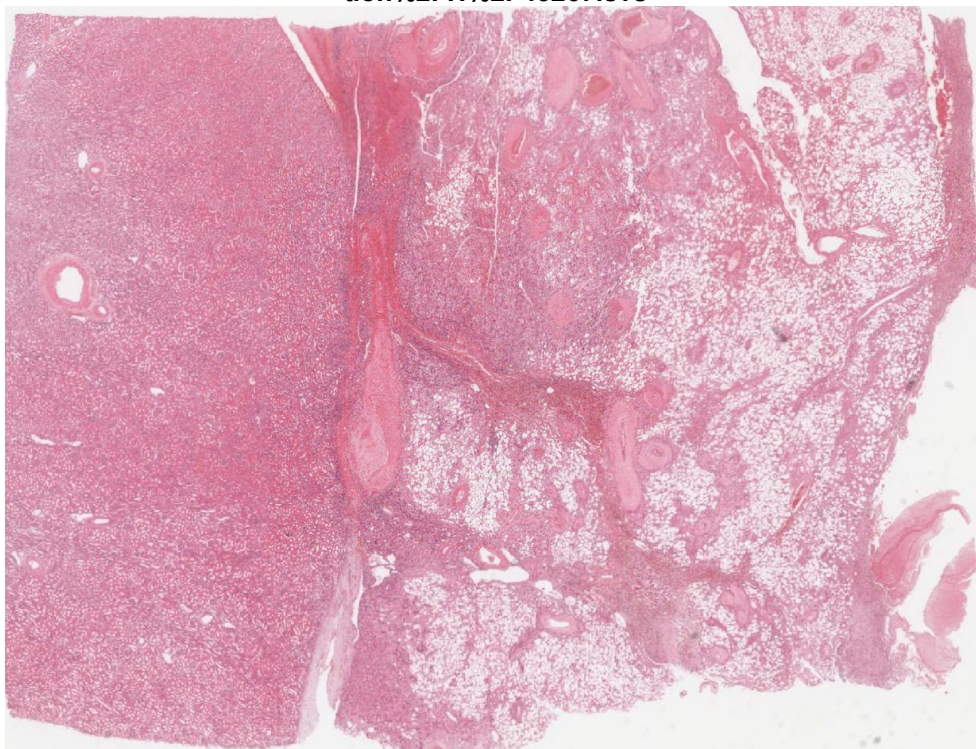
Fonte: Aatoria própria (2020)

Figura 59 – Apresenta a lâmina completa 33836 utilizada para a avaliação de desempenho da rede, tamanho em pixel 98880x66650, https://www.virtualpathology.leeds.ac.uk/slides/browser/view_flash.php?path=%2FResearch_4%2FSlide_Library%2FR_Bishop_Collection%2FCard_index_Set%2FRenal%2F33836.svs



Fonte: Aatoria própria (2020)

Figura 60 – Apresenta a lâmina completa 43267 utilizada para a avaliação de desempenho da rede, tamanho em pixel 98880x66650, https://www.virtualpathology.leeds.ac.uk/slides/browser/view_flash.php?path=%2FResearch_4%2FSlide_Library%2FRenal_EQA_Collection%2FH%2F43267.svs



Fonte: Autoria própria (2020)

APÊNDICE D - Códigos

O Código 1, em linguagem Python, mostra o programa implementado para baixar as imagens do site Bio-Atlas.

Código 1 – Lógica desenvolvida para fazer o download das imagens no site Bio-Atlas

```

1.  #-*- coding: utf-8 -*-
2.  import urllib.request
3.  import time
4.
5.
6.  #Número da imagem que está sendo baixada
7.  v_s = 1292
8.  #Tamanho da imagem 0 "maior Zoom"
9.  v_z = 0
10. #Número total de imagens
11. v_finish1 = 9900
12. #Número da primeira imagem
13. v_start1 = 0
14.
15. #Função para download da imagem
16. def getImagem(v_start, v_finish):
17.
18.     for y in range(v_start, v_finish):
19.         print("Y", y)
20.         #Url do site que estamos buscando a imagem
21.         url='http://bio-
    atlas.psu.edu/tile.jpeg.php?s='+str(v_s)+'&z='+str(v_z)+'&i='+str(y)
22.         #Caminho onde a imagem será salva
23.         img_name = './teste/Imagem_1309/lmg_Exp1309_1_'+ str(y) +'.jpg'
24.
25.         try:
26.             urllib.request.urlretrieve(url,img_name)
27.             time.sleep(0) # Sleep for 10 seconds
28.         except:
29.             #Caso ocorra um erro no download, uma exceção é lançada,
30.             #chamando de forma recursiva a função getImagem,
31.             #garantindo que todas as imagens serão baixadas
32.             print("Ocorreu exceção")
33.             getImagem(y, v_finish)
34.
35.     print("Fim do processamento")
36.
37.
38. #Processamento inicial
39. getImagem(v_start1, v_finish1)

```

Fonte: Autoria própria (2020)

O Código 2 foi desenvolvido para montar as imagens baixadas do site Bio-Atlas no tamanho padrão de 2048x2048 (i.e., tamanho padrão definido para apresentar para a rede) este código pega uma imagem de qualquer tamanho (i.e., no nosso caso 256x256) e reconstrói a mesma para o tamanho definido como padrão.

Código 2 – Programa em Python para montar as imagens no tamanho de 2048x2048.

```

1. import os
2. import numpy as np
3. import cv2
4.
5. # from PIL import Image, ImageFilter
6. w_tile, h_tile = (256, 256) # typical
7. multiples = 4
8.
9. image_factor_w = 2 * multiples # if = 2 there will be 8 width tiles, if 3 there will
   be 12 width tiles
10. image_factor_h = 2 * multiples # if = 2 there will be 6 high tiles, if 3 there will
    be 9
11. # This is teh size of the output new bigger TILES for latter Deep Learning pro-
    cessing
12. w_imageOut = w_tile * image_factor_w
13. h_imageOut = h_tile * image_factor_h
14. print(w_imageOut, h_imageOut )
15. # This is the Original size of the Whole Slide Image (example = 33000 x
    24000)
16. #w_Orig_Entire_image, h_Orig_Entire_image = ()
17. #div_h = h_Orig_Entire_image / h_tile
18. #div_w = w_Orig_Entire_image / w_tile
19. # This is the information on the number of X and Y tiles
20. div_w = 76 #113 # confirm and edit HERE after downloading images
21. #div_h = 63 #92 # confirm and edit HERE after downloading images
22.
23. #717 76 66
24. # 867 84 38
25.
26.
27. # define input folder for small tile images
28. input_dir = './teste/Imagem_569'
29. aaa = os.listdir(input_dir)
30. number_files = len(aaa)
31.
32.
33. div_h = number_files / div_w #Get the number of lines of the images
34. div_h = int(div_h)
35. print(aaa)
36. # create output folder for bigger tile images
37. output_dir = input_dir + '_out/'
38. print(output_dir)

```

```

39. if not os.path.exists(output_dir):
40.     os.makedirs(output_dir)
41. # define typical filename to be used for input images
42. typical_input_img_name = 'Img_Exp569_1_'
43. # assuming numbered images prepare to read the first one
44. i_th_image = 0 # the very first index for the input tiles
45. big_tile_number = 1
46. for y in range(0, div_h, image_factor_h):
47.     for x in range(0, div_w, image_factor_w):
48.         # Need for controlling the new big tiles formation
49.         ###output_image_4_3_factor = np.zeros((h_imageOut, w_imageOut,4),
dtype=np.uint8)
50.         # this might be 8x6 or 12x9
51.         #y_internal_value = 8 # might be 12
52.         #x_internal_value = 6 # might be 9
53.         #image_factor_h2 = image_factor_h
54.         if(div_h - y < image_factor_h):
55.             image_factor_h2 = div_h - y
56.         else:
57.             image_factor_h2 = image_factor_h
58.         print(y, image_factor_h2)
59.
60.
61.         if(div_w - x < image_factor_w):
62.             image_factor_w2 = div_w - x
63.         else:
64.             image_factor_w2 = image_factor_w
65.         print(x, image_factor_w2)
66.
67.         for y1 in range(image_factor_h2):
68.             for x1 in range(image_factor_w2):
69.                 i_th_image = (y+y1)*div_w + x+x1
70.                 input_name_ith = input_dir + '\\' + typical_input_img_name +
str(i_th_image) + '.jpg'
71.                 print(input_name_ith)
72.
73.                 #print(i_th_image)
74.                 img = cv2.imread(input_name_ith)
75.                 # Converte a imagem para RGB opencv
76.                 #r, g, b = cv2.split(img)
77.                 #img_in_rgb = cv2.merge((b, g, r))
78.                 img_in_rgb = img
79.                 if(x1==0):
80.                     matrix_out_v = img_in_rgb
81.                 else:
82.                     matrix_out_v = np.hstack((matrix_out_v,img_in_rgb))
83.                 #output_image_4_3_factor() = img_in_rgb
84.
85.                 if(y1==0):
86.                     matrix_out_h = matrix_out_v

```



```

87.     else:
88.         matrix_out_h = np.vstack((matrix_out_h, matrix_out_v))
89.         big_tile_number = big_tile_number + 1
90.         print('preparing new Big Tile', big_tile_number)
91.         output_name_ith = output_dir + typical_input_img_name +
           str(big_tile_number) + '.jpg'
92.         print(output_name_ith)
93.         output_image_4_3_factor = matrix_out_h
94.         cv2.imwrite(output_name_ith, output_image_4_3_factor)

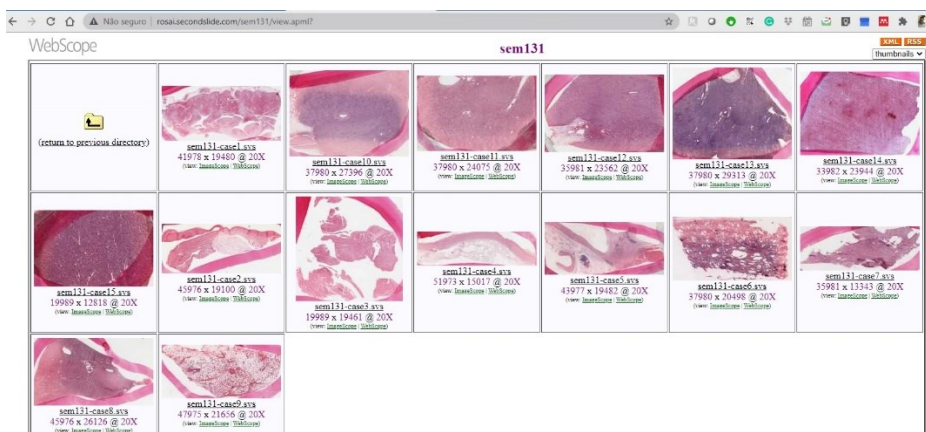
```

Fonte: Autoria própria (2020)

Para baixar as imagens das bases de imagens Rosai's e Leeds, é necessário entrar no site, baixar os arquivos (e.g., 1111.sis), abrir na máquina local com o software Aperio ImageScope e através da função de download, baixar o arquivo na máquina local. A figura 61 demonstra o passo a passo para fazer o download da imagem, onde selecionamos a imagem, com base na tag gerada no site, para gerar o arquivo com extensão sis, este arquivo é compatível com o Aperio ImageScope, que abre a imagem nos proporcionando a opção para que possamos efetuar o download.

Figura 61 – Site Rosai's para *download* da imagem, (a) mostra o site com as imagens que podem ser baixadas, no nosso exemplo estamos usando o case10.svs, selecionamos ImageScope, (b) após a seleção é aberta a tela com as tags para abrir a imagem no Aperio ImageScope, devemos criar um arquivo .sis e salvar a tag, (c) abrindo o arquivo .sis no software Aperio ImageScope, (d) para efetuar o download selecionamos *extract region* e selecionamos a parte que queremos extrair, ou podemos selecionar a imagem inteira.

(a)

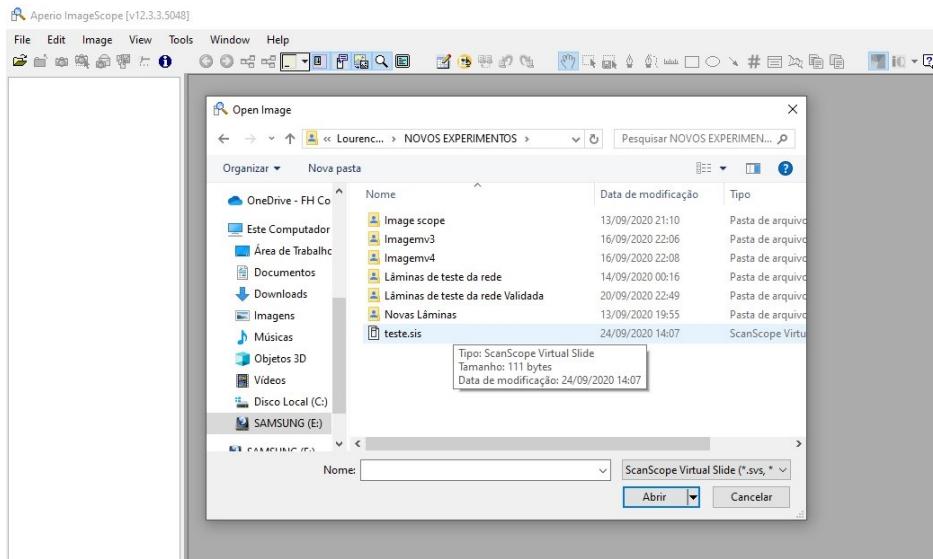


(b)

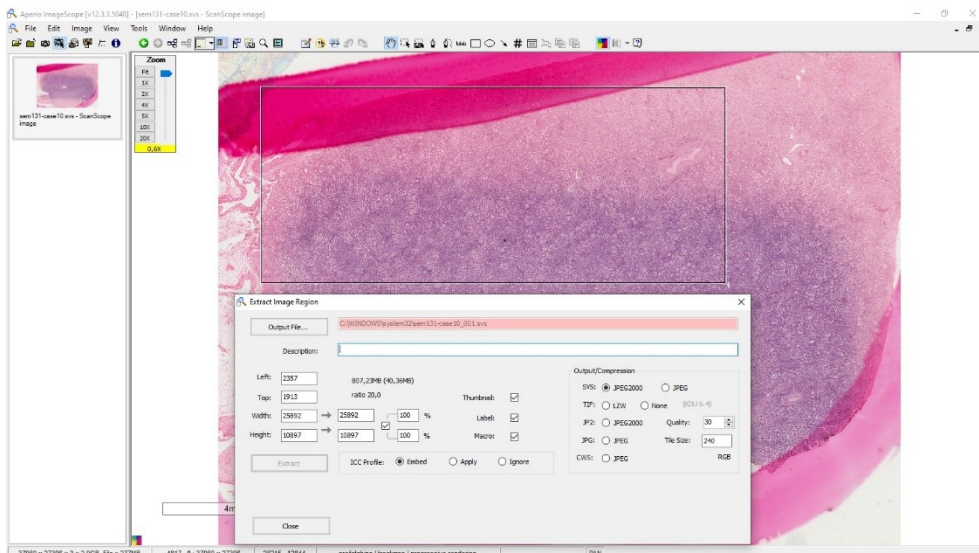
```

rosai.seconslide.com/sem131/sem131-case10.svs/is.sis?chost=rosai.seconslide...
<SIS version="1.0">
  <Image>
    <URL>//rosai.seconslide.com/sem131/sem131-case10.svs</URL>
  </Image>
</SIS>
    
```

(c)



(d)



Fonte: Autoria própria (2020)

Após todos os arquivos baixados localmente, e convertidos para a extensão jpeg, foi realizada o recorte das imagens para o tamanho padrão (i.e., 2048x2048),

desenvolvemos um programa em Python para quebrar as imagens em várias imagens de 2048x2048. O código 3 mostra a lógica utilizada para processar as imagens e dividi-las em outras imagens com o tamanho padrão.

Código 3 – Código para dimensionamento das imagens das bases de dados de imagens em imagens de 2048x2048.

```

1. # -*- coding: utf-8 -*-
2. """
3. Created on Tue Dec 3 22:42:15 2019
4.
5. @author: lbarbosa
6. """
7.
8. import openslide as os
9. import numpy as np
10. import cv2
11. #from PIL import Image, ImageFilter
12.
13. vl_level = 0
14.
15. # imagem
16. openSlideObj = os.OpenSlide('./Experimento 3/Ima-
    gen_SVS/Imag_exp_L/img_exp_L_2/131-case15_Baixo.svs')
17.
18.
19. print(openSlideObj.level_count)
20. print(openSlideObj.level_dimensions[vl_level])
21.
22. W, H = openSlideObj.level_dimensions[vl_level]
23. w, h = (2480, 2480)
24.
25. Sobreposicao_h = 0
26. Sobreposicao_w = 0
27.
28. h_dist = h - Sobreposicao_h
29. w_dist = w - Sobreposicao_w
30.
31.
32. div_h = H / h_dist
33. div_w = W / w_dist
34.
35. div_h = int(div_h)
36. div_w = int(div_w)
37. contador = 0
38.
39. for y in range(div_h):
40.     for x in range(div_w):
41.         print("Y", y*h_dist)

```

```
42.     print("X", x*w_dist)
43.
44.     contador = contador + 1
45.
46. # Recorta a imagem.
47. #imagem2 = openSlideObj.read_region(location(x,y),level,size(width, height))
48.     imagem2 = openSlideObj.read_region((x*w_dist,
    y*h_dist),vl_level,(int(w), int(h)))
49. #     imagem2 = openSlideObj.read_region((0, 0),vl_level,(W, H))
50.     imagem3 = np.array(imagem2,dtype = np.uint8)
51.
52. #Converte a imagem para RGB padrão opencv
53.     r,g,b,a = cv2.split(imagem3)
54.     imagem4 = cv2.merge((b,g,r))
55.
56. #salva imagem no diretório
57.     cv2.imwrite('./Experimento/Img_ExpL_2_'+ str(contador) +'.jpg', ima-
    gem4)
58.
59.
60. print("Fim do processamento")
61. cv2.waitKey(0)
```

Fonte: Autoria própria (2020)