
**Segmentação Automática de Canais Ósseos em
Imagens Histológicas Utilizando Redes Neurais
Completamente Convolucionais**

Igor Gonçalves Ribeiro Silva



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia
2024

Igor Gonçalves Ribeiro Silva

**Segmentação Automática de Canais Ósseos em
Imagens Histológicas Utilizando Redes Neurais
Completamente Convolucionais**

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Bruno Augusto Nassif Travençolo

Uberlândia

2024

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

S586s
2024 Silva, Igor Gonçalves Ribeiro, 1996-
Segmentação automática de canais ósseos em imagens histológicas
utilizando redes neurais completamente convolucionais [recurso
eletrônico] / Igor Gonçalves Ribeiro Silva. - 2024.

Orientador: Bruno Augusto Nassif Travençolo.
Dissertação (Mestrado) - Universidade Federal de Uberlândia,
Programa de Pós-graduação em Ciência da Computação.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.di.2024.5120>
Inclui bibliografia.
Inclui ilustrações.

1. Computação. I. Travençolo, Bruno Augusto Nassif, 1981-,
(Orient.). II. Universidade Federal de Uberlândia. Programa de Pós-
graduação em Ciência da Computação. III. Título.

CDU: 681.3

André Carlos Francisco
Bibliotecário Documentalista - CRB-6/3408



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Ciência da
Computação

Av. João Naves de Ávila, 2121, Bloco 1A, Sala 243 - Bairro Santa Mônica, Uberlândia-MG,
CEP 38400-902

Telefone: (34) 3239-4470 - www.ppgco.facom.ufu.br - cpgfacom@ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação				
Defesa de:	Dissertação, 27/2024, PPGCO				
Data:	14 de agosto de 2024	Hora de início:	14:00	Hora de encerramento:	16:00
Matrícula do Discente:	12212CCP010				
Nome do Discente:	Igor Gonçalves Ribeiro Silva				
Título do Trabalho:	Segmentação Automática de Canais Ósseos em Imagens Histológicas Utilizando Redes Neurais Completamente Convolucionais				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Ciência de Dados				
Projeto de Pesquisa de vinculação:	Processo: 306436/2022-1 PQ-CNPq Título: Da Visão Computacional à Visualização: novas proposições para pesquisa em eScience				

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Marcelo Zanchetta do Nascimento - FACOM/UFU, Thaína Aparecida Azevedo Tosta- UNIFESP e Bruno Augusto Nassif Travençolo - FACOM/UFU, orientador do candidato.

Os examinadores participaram desde as seguintes localidades: Thaína Aparecida Azevedo Tosta - São José dos Campos/SP. Os outros membros da banca e o aluno participaram da cidade de Uberlândia.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Bruno Augusto Nassif Travençolo, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir ao candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

Aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Thaína Aparecida Azevedo Tosta, Usuário Externo**, em 19/08/2024, às 11:28, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Bruno Augusto Nassif Travençolo, Professor(a) do Magistério Superior**, em 19/08/2024, às 14:02, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marcelo Zanchetta do Nascimento, Professor(a) do Magistério Superior**, em 19/08/2024, às 16:49, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5587073** e o código CRC **E0C9A40E**.

Referência: Processo nº 23117.050255/2024-97

SEI nº 5587073

Agradecimentos

Agradeço à Universidade Federal de Uberlândia, em especial a Faculdade de Computação, todo o seu corpo docente, direção e administração por esta oportunidade de crescimento pessoal, profissional e acadêmico.

Ao meu orientador Bruno Augusto Nassif Travençolo por todo o suporte e incentivo e compreensão presentes nas incontáveis aulas e reuniões e todas as suas correções e ensinamentos que muito contribuíram para minha formação como pessoa e pesquisador.

À professora Paula Dechichi pela disponibilização do material biológico utilizado no desenvolvimento do trabalho, bem como o incentivo e apoio para o desenvolvimento dele.

A minha esposa Camila que, além da participação ativa no desenvolvimento deste trabalho, pelo incentivo em retornar à academia e todo o apoio, amor e compreensão durante esse período. Sem ela, por diversos motivos, este trabalho não seria realizado.

A minha mãe Maria de Fátima, que nunca mediu esforços nem incentivos para que eu pudesse construir uma vida digna através dos estudos.

A minha tia Geralda Gislene deu todo o apoio e carinho durante minha graduação, tornando possível tal importante passo em minha carreira.

Resumo

Neste trabalho é proposto um método de segmentação de canais ósseos em imagens histológicas de lâmina inteira. O método faz uso de uma rede neural originalmente desenvolvida com o intuito de segmentar tumores derivados da cavidade oral em imagens histológicas coradas com Hematoxilina e Eosina, e foi adaptado para o contexto de canais ósseos. O conjunto de dados é composto por 65 imagens de lâmina inteira extraídas a partir do fêmur de ratos saudáveis da linhagem Wistar e coradas com Hematoxilina e Eosina. Com a ajuda de um especialista em histologia, as imagens foram analisadas e seus canais ósseos foram manualmente marcados gerando um conjunto de máscaras binárias. As imagens de ambos os conjuntos (imagens originais e binárias) foram então divididas em sub imagens de 640×640 pixels. A rede foi treinada e validada com 2037 sub imagens. O treinamento contou ainda com uma estratégia de aumento de dados com sete possíveis variações das imagens. O método foi avaliado comparando-se as regiões segmentadas pela rede com as marcações do especialista. Foram calculadas a acurácia, especificidade, sensibilidade, precisão, Intersecção sobre União e *f1-score* das segmentações resultantes. Além disso, foi feita uma comparação com outro método de segmentação automática de canais ósseos da literatura. O método validado por este trabalho mostrou-se eficiente e superior ao método com o qual foi comparado, apresentando *f1-score* de 84,9% e Intersecção sobre União de 73,7%, além de bons resultados qualitativos.

Palavras-chave: Segmentação de imagens. Canais Ósseos. Redes Neurais Convolucionais.

Abstract

In this work, a method for segmenting bone canals in whole-slide histological images is proposed. The method utilizes a neural network originally developed to segment tumors derived from the oral cavity in Hematoxylin and Eosin-stained histological images, and it has been adapted for the context of bone canals. The dataset consists of 65 whole-slide images stained with Hematoxylin and Eosin, extracted from the femur of healthy Wistar rats. With the assistance of a histology expert, the images were analyzed, and their bone canals were manually marked, generating a set of binary masks. Images from both sets (original and binary images) were then divided into 640×640 pixel-sized sub-images. The network was trained and validated with 2037 sub-images. Training also included a data augmentation strategy with seven possible image variations. The method was evaluated by comparing the regions segmented by the network with the specialist's annotations. Accuracy, specificity, sensitivity, precision, Intersection over Union, and f1-score of the resulting segmentations were calculated. Additionally, a comparison was made with another automatic bone canal segmentation method of the literature. The method validated by this work proved to be efficient and superior to the compared method, with an f1-score of 84.9% and Intersection over Union of 73.7%, along with good qualitative results.

Keywords: Image Segmentation. Bone Canals. Convolutional Neural Networks.

Lista de ilustrações

Figura 1 – Estrutura básica de um osso longo.	19
Figura 2 – Corte transversal da diáfise de um osso longo	20
Figura 3 – Estrutura básica de uma rede neural convolucional.	23
Figura 4 – Exemplo de convolução entre duas matrizes bidimensionais.	24
Figura 5 – Agrupamento utilizando uma função de <i>max-pooling</i>	24
Figura 6 – Estrutura básica da rede U-Net.	26
Figura 7 – Exemplo de entrada e saída do método proposto por (GONDIM et al., 2021)	29
Figura 8 – Exemplo de marcações feitas pelo especialista em (GONDIM et al., 2021)	29
Figura 9 – Exemplos de saídas do método proposto em (JULIA et al., 2021).	30
Figura 10 – Principais elementos gráficos presentes no método proposto por (SANTOS et al., 2022).	31
Figura 11 – Diagrama do método proposto.	33
Figura 12 – Exemplo de imagem utilizada no método proposto.	34
Figura 13 – Marcação do especialista para o método proposto.	35
Figura 14 – Arquitetura da rede utilizada	35
Figura 15 – Região de imagem marcada manualmente e sua respectiva máscara binária.	36
Figura 16 – Transformações utilizadas no aumento de dados. Fonte: Autor.	37
Figura 17 – Curvas de acurácia e perda ao longo do treinamento.	40
Figura 18 – Saída da rede neural.	40
Figura 19 – Métricas obtidas na análise por pixel de ambos os treinamentos.	41
Figura 20 – Comparação entre marcação manual feita por especialista e saída do método.	43
Figura 21 – Imagem marcada pela rede com regiões ampliadas	44
Figura 22 – Exemplos de componentes conectados obtidos pelo método proposto.	45
Figura 23 – Métricas obtidas na análise por canal para ambos os treinamentos.	46

Figura 24 – Comparação entre marcação manual feita por especialista e saída do método.	48
Figura 25 – Comparação de regiões ampliadas das marcações manuais feita por especialista e de saídas do método	49
Figura 26 – Exemplos de componentes conectados obtidos pelo método proposto por (GONDIM et al., 2021).	50
Figura 27 – Exemplo de falha do algoritmo proposto por (GONDIM et al., 2021). .	51

Lista de tabelas

Tabela 1 – Médias de acurácia, precisão, <i>f1-score</i> , sensibilidade e especificidade para cada <i>threshold p</i> testado na análise por pixel.	42
Tabela 2 – Médias de precisão, <i>f1-score</i> , sensibilidade e IoU para cada limiar <i>p</i> testado na análise por canal.	47
Tabela 3 – Médias de precisão, <i>f1-score</i> , sensibilidade e Intersecção sobre União para cada um dos métodos de segmentação testados.	48

Lista de abreviaturas

RNC Redes Neurais Convolucionais

RNCC Redes Neurais Completamente Convolucionais

HE Hematoxilina e Eosina

TTA Treinamento por Transferência de Aprendizado

TZ Treinamento a partir do zero

Sumário

1	INTRODUÇÃO	14
1.1	Motivação	15
1.2	Objetivo	16
1.3	Hipótese	16
1.4	Contribuições	17
1.5	Organização do trabalho	17
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	O tecido ósseo e a rede de canais ósseos	18
2.2	Aprendizado de máquina	20
2.3	Aprendizado Profundo	21
2.4	Transferência de aprendizado	21
2.5	Redes Neurais Convolucionais	22
2.5.1	Camadas convolucionais	22
2.5.2	Camadas de agrupamento	24
2.5.3	Camadas de achatamento	24
2.5.4	Camada densa ou conectada	25
2.6	U-Net	25
2.7	Métricas de validação	26
2.7.1	Acurácia	26
2.7.2	Especificidade	27
2.7.3	Precisão	27
2.7.4	Sensibilidade	27
2.7.5	<i>F1-Score</i>	27
2.7.6	Intersecção sobre União	28
2.8	Trabalhos Relacionados	28

3	MATERIAIS E MÉTODOS	32
3.1	Conjunto de dados	32
3.2	Arquitetura da Rede Neural	34
3.3	Método	36
4	RESULTADOS E DISCUSSÃO	39
4.1	Treinamento da Rede	39
4.2	Análise por pixel	39
4.3	Análise por canal	42
4.3.1	Comparação	45
5	CONCLUSÃO	52
5.1	Contribuições em Produções Bibliográficas	53
	REFERÊNCIAS	54

Introdução

A histologia é a área da biologia que estuda tecidos e composição de órgãos. Ela possui grande relevância na área acadêmica pois tem como proposta estudar a estrutura e a função dos tecidos, realizar análises quantitativas, avaliar respostas a tratamentos e investigar aspectos do desenvolvimento embrionário e da fisiologia. Outras aplicações importantes da histologia são encontradas na medicina, em que é empregada em diagnósticos médicos, auxiliando patologistas a identificar e caracterizar doenças, tumores, infecções e outras alterações nos tecidos (JUNQUEIRA; CARNEIRO, 1985).

As análises histológicas são comumente feitas a partir de imagens denominadas cortes histológicos, que podem ser obtidas em laboratório por meio de um longo processo de preparo. A análise dessas imagens de lâmina inteira muitas vezes são demoradas e demandam grande esforço de um especialista (LINHARES et al., 2022). Dada a importância da análise histológica, é interessante desenvolver métodos que possam facilitá-la, tornando-a mais rápida e acessível para pesquisadores e profissionais da saúde.

O tecido ósseo é um tecido conjuntivo que tem grande relevância em pesquisas da área da histologia devido à sua estrutura, composição e propriedades regenerativas. Desenvolver formas de automatizar a análise de imagens histológicas de tecido ósseo pode auxiliar estudos relacionados à estrutura do tecido, sua regeneração e efeitos de tratamentos de doenças (LINHARES et al., 2019). Porém é um desafio automatizar tais análises devido a vários fatores como o tamanho das imagens, complexidade e quantidade de estruturas, irregularidades e rasgos no tecido decorrentes do processo de preparo do corte histológico (GONDIM et al., 2021).

Nos últimos anos, o aprendizado de máquina tem ganhado destaque tanto no meio acadêmico como no corporativo por apresentar resultados satisfatórios na execução de tarefas complexas, inclusive em relação ao processamento de imagens (MIKLOSIK; EVANS, 2020). Uma das abordagens do aprendizado de máquina é o aprendizado supervisionado, cujo princípio é o uso de uma base de dados para treinar um modelo de forma que o mesmo aprenda a realizar uma determinada tarefa. Esse conjunto de dados deve estar organizado de forma a conter exemplos diversos de entradas de dados e suas respectivas

saídas esperadas (MONARD; BARANAUSKAS, 2003).

Ao se trabalhar com modelos de aprendizado de máquina, o conjunto de dados a ser utilizado pode ser um fator limitante, pois é preciso um conjunto bem estruturado para realizar a tarefa desejada para que o treinamento não seja enviesado e apresente bons resultados (PAULLADA et al., 2021). Dessa forma, para criar e treinar um modelo de aprendizado de máquina para a execução de uma tarefa, muitas vezes é necessário um árduo trabalho prévio de elaboração de um conjunto de dados que viabilize o treinamento.

Conjuntos de imagens são amplamente utilizados em técnicas de aprendizado profundo (do inglês, *deep learning*), uma subárea do aprendizado de máquina. O uso de redes neurais, modelos de aprendizado de máquina que tentam simular estruturas do cérebro humano, tem apresentado resultados interessantes no campo da visão computacional, especialmente as Redes Neurais Convolucionais (RNC). Tais redes se baseiam na forma como os seres humanos percebem e aprendem características chave das imagens (RAWAT; WANG, 2017).

O uso do aprendizado profundo na área médica vem ganhando força e importância especialmente desde o surgimento das redes do tipo U-Net, arquitetura proposta em 2015 por (RONNEBERGER; FISCHER; BROX, 2015), que vem sendo amplamente utilizada para segmentação de imagens biomédicas graças à sua capacidade de realizar classificação a nível de pixel. A utilização de tais técnicas pode ser de grande ajuda para especialistas em análises clínicas, diagnósticos e pesquisas, tornando o processo mais ágil (ESTEVA et al., 2021).

1.1 Motivação

Sabendo da importância da visão computacional nas ciências biomédicas, e do importante papel que o aprendizado profundo vem desempenhando na área de visão computacional, é interessante oferecer ferramentas que façam algum tipo de processamento automático em imagens histológicas, tais como classificação de imagens ou segmentação semântica.

Existem alguns métodos que visam resolver problemas específicos no processamento automático de imagens, como segmentação de alguma estrutura específica a partir da aplicação de determinados procedimentos nas imagens, como o proposto em (GONDIM et al., 2021). Porém, tais métodos muitas vezes não apresentam invariância, ou seja, não apresentam robustez a variações nas entradas, como tamanho das imagens, ruídos, variações de cor e posicionamento. Isso dificulta a utilização dessas ferramentas, visto que a parametrização ideal pode ser diferente para cada entrada (LINHARES et al., 2022). Esse cenário definiu a hipótese de se criar uma ferramenta de processamento de imagens que possa se ajustar automaticamente para a realização de uma determinada tarefa independentemente das características específicas de cada imagem a ser analisada.

Para a realização desse tipo de trabalho, as RNCs têm apresentado ótimos resulta-

dos, pois por meio da convolução aprendem padrões de imagens em um nível local e os identificam na imagem independente de posição, tamanho ou rotação (MUELLER; MASSARON, 2019). Portanto torna-se possível utilizar uma RNC que realize segmentação de estruturas de interesse em imagens histológicas de tecido ósseo com desempenho e precisão satisfatórias.

Como já mencionado, utilizar uma rede neural para este tipo de tarefa requer um conjunto de dados estruturado de forma adequada. Existem vários conjuntos de dados de imagens médicas disponíveis para uso, por exemplo o ALL-IDB (Labati; Piuri; Scotti, 2011) e ErythrocytesIDB (GONZÁLEZ-HIDALGO et al., 2015), que são conjuntos de amostras de sangue; e o COVID-19 *Radiography Database* (CHOWDHURY et al., 2020), conjunto de imagens de raio-x de pulmões de pacientes com COVID-19. Apesar de existirem vários conjuntos de imagens médicas, como os mencionados acima, a escassez de conjuntos de dados adequados para tarefas específicas é comumente um desafio ao se trabalhar com treinamento de RNCs na área médica. Muitas vezes se faz necessário o desenvolvimento de um conjunto de imagens próprio, criado especificamente para a realização da uma tarefa (KIRYATI; LANDAU, 2021).

No caso da segmentação de canais ósseos em imagens histológicas observou-se uma escassez de dados destinado especificamente a tal tarefa, o que se tornou um desafio para a realização deste trabalho e levou à necessidade de se criar um conjunto de imagens próprio para tal. Porém o processo de criação de um conjunto de dados pode ser oneroso e muitas vezes depende da disponibilidade de um especialista da área em questão (GUPTA et al., 2020).

1.2 Objetivo

Partindo da ideia apresentada, o objetivo deste trabalho é realizar uma investigação de modelo RNC para segmentação de canais ósseos em imagens histológicas de tecido ósseo. Quanto aos objetivos específicos, o primeiro é estabelecer critérios de treinamento para se treinar uma RNC para realizar segmentação de canais ósseos em imagens histológicas. Um segundo objetivo específico deste trabalho é a criação e disponibilização de um conjunto de imagens apropriado para a tarefa em questão.

Um desafio deste trabalho foi a preparação da base de imagens, que foi feita por meio da marcação da região de interesse em várias imagens de lâmina inteira. Tais marcações foram realizadas com a ajuda de um especialista em histologia, que treinou os autores, supervisionou e validou as marcações realizadas.

1.3 Hipótese

Este trabalho busca validar a seguinte hipótese:

É possível utilizar Redes Neurais Convolucionais para realizar a segmentação de canais ósseos em imagens histológicas de lâmina inteira coradas com Hematoxilina e Eosina (HE).

1.4 Contribuições

A principal contribuição deste trabalho é demonstrar a viabilidade do uso de uma RNC para a tarefa de segmentação de canais ósseos em imagens histológicas, permitindo o uso de RNCs em outras tarefas da histologia relacionadas a imagens de tecido ósseo e para o desenvolvimento de novas tecnologias que possam contribuir para pesquisas dessa área.

Outra contribuição é a disponibilização pública de um conjunto de imagens histológicas de tecido ósseo, que pode ser útil para outros pesquisadores na elaboração e desenvolvimento de suas pesquisas. Tal conjunto pode ser evoluído em trabalhos posteriores para que seja possível realizar outras tarefas de visão computacional em imagens de tecido ósseo, como classificação ou segmentação de outras estruturas presentes nesse tipo de tecido.

1.5 Organização do trabalho

Este documento está organizado da seguinte forma:

- ❑ O Capítulo 2 traz a fundamentação teórica referente ao contexto do trabalho e às técnicas empregadas no desenvolvimento do método. Também serão apresentadas e justificadas as métricas utilizadas para a validação do método. O capítulo apresenta ainda alguns trabalhos que apresentam correlação com o estudo desenvolvido.
- ❑ O Capítulo 3 apresenta o conjunto de dados e descreve o processo utilizado para sua elaboração. Também é descrita a arquitetura da rede neural utilizada e a metodologia aplicada para seu treinamento e validação.
- ❑ O Capítulo 4 traz os resultados obtidos e discute sua significância, além de comparar os resultados com os de outro método encontrado na literatura.
- ❑ O Capítulo 5, por fim, conclui o trabalho e discute possíveis pontos de evolução.

Fundamentação Teórica

Neste capítulo são abordados os conceitos fundamentais que embasam o desenvolvimento desta pesquisa e os principais trabalhos correlatos.

2.1 O tecido ósseo e a rede de canais ósseos

O tecido ósseo é um dos objetos de estudo da histologia e chama a atenção de pesquisadores e profissionais da área médica especialmente devido às suas propriedades regenerativas. Macroscopicamente os ossos longos podem ser divididos em duas regiões: diáfise e epífise. As epífises são as extremidades do osso e são compostas por osso esponjoso, ou seja, com muitas cavidades intercomunicantes (JUNQUEIRA; CARNEIRO, 1985). Já a diáfise é a região intermediária do osso, sendo mais fina e apresentando osso compacto – sem cavidades – em sua maior parte (JUNQUEIRA; CARNEIRO, 1985). Na Figura 1(a) são mostradas tais regiões e suas respectivas composições.

A maior parte da diáfise é composta por túneis longitudinais formados por lamelas concêntricas. Tais túneis são conhecidos como sistema de Havers e formam em seus centros canais que são percorridos por vasos sanguíneos, linfáticos e nervos. Também existem canais transversais que conectam canais de Havers adjacentes chamados canais de Volkmann, formando assim a rede de canais ósseos (JUNQUEIRA; CARNEIRO, 1985). Na Figura 1(b) são mostradas as estruturas mencionadas destacando os canais de Havers e de Volkmann.

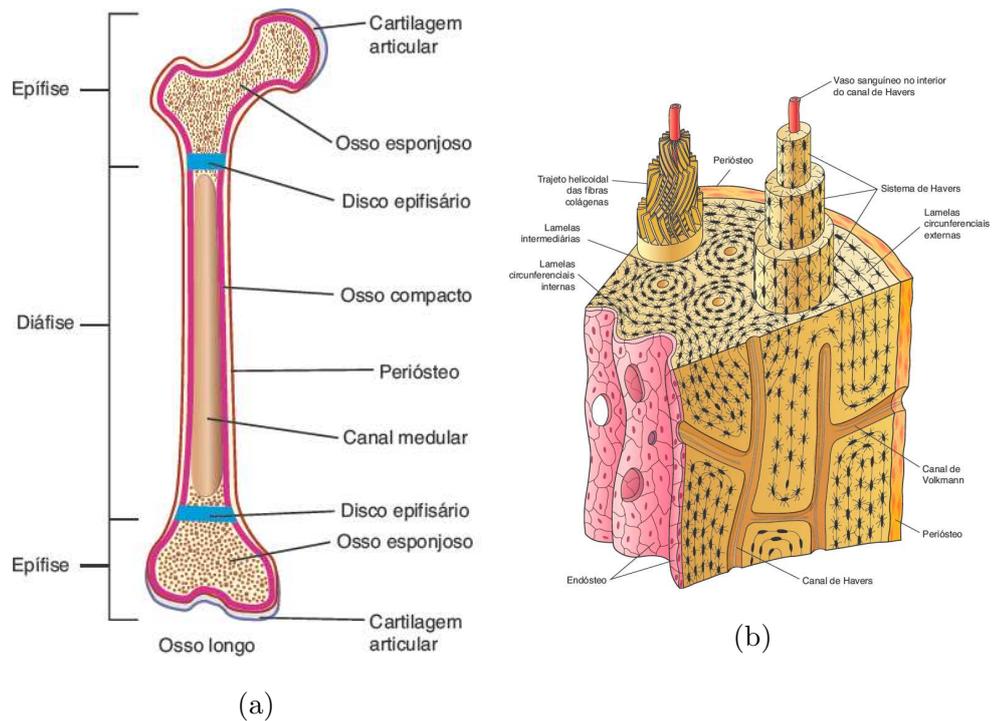


Figura 1 – Estrutura básica de um osso longo. Em (a) os principais componentes anatômicos de um osso longo. Em (b) os canais de Havers e Volkmann, que compõem o osso compacto, principal componente da diáfise, região intermediária de ossos longos. Fonte: (JUNQUEIRA; CARNEIRO, 1985)

Os vasos sanguíneos e linfáticos que percorrem os canais nutrem os osteócitos, que são estruturas que cumprem um importante papel na manutenção e integridade da matriz óssea. Dessa forma a rede de canais ósseos cumpre um importante papel de nutrição do tecido contribuindo para seu desenvolvimento e no processo de reparo ósseo (JUNQUEIRA; CARNEIRO, 1985).

Na Figura 2 é mostrado um corte transversal na diáfise de um osso longo e um exemplo de imagem histológica dessa região, destacando os canais ósseos e os osteócitos.

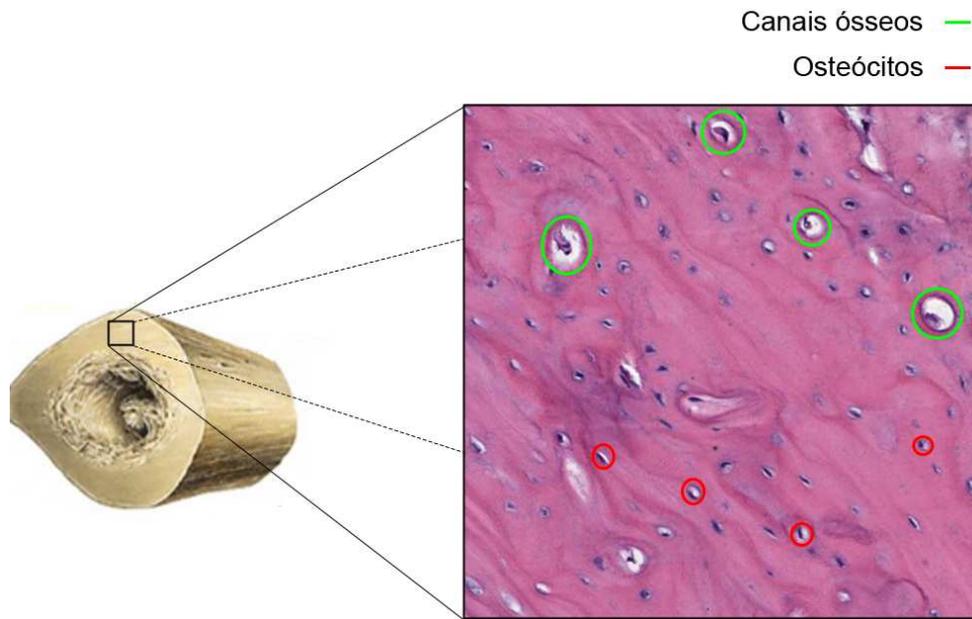


Figura 2 – Corte transversal da diáfise de um osso longo com uma região ampliada destacando alguns canais ósseos e osteócitos. Fonte: Autor.

2.2 Aprendizado de máquina

Em 1959, Arthur Lee Samuel definiu o aprendizado de máquina como “o campo de estudo que dá aos computadores a habilidade de aprender sem serem explicitamente programados” (SIMON, 2013). De fato o aprendizado de máquinas automatizado permite, por exemplo, que computadores aprendam a realizar tarefas analisando uma base de dados previamente rotulada, além de aprimorar seu aprendizado à medida que novos dados lhe são apresentados (MONARD; BARANAUSKAS, 2003).

O aprendizado de máquina é um área da inteligência artificial que se apoia em técnicas e conceitos da matemática e estatística. Esse aprendizado é dividido em quatro tipos: aprendizado supervisionado, não supervisionado, auto supervisionado e por reforço. Todos possuem uma variedades de algoritmos, sendo que, para cada problema a ser resolvido, um determinado algoritmo pode ser mais adequado que os demais (MUELLER; MASSARON, 2019).

Como mencionado acima, uma das técnicas de aprendizado de máquina é o aprendizado supervisionado. Nessa abordagem utiliza-se uma base de dados na qual para cada entrada o resultado desejado é conhecido. As entradas com seus respectivos resultados esperados, também chamados de rótulos, são apresentados ao algoritmo em uma etapa conhecida como etapa de treinamento, em que o algoritmo ajusta, com base nos dados fornecidos, uma série de parâmetros internos a fim de encontrar padrões que o permitam estimar resultados para novas entradas (MONARD; BARANAUSKAS, 2003). A saída

pode ser qualitativa, tratando-se assim de uma tarefa de classificação; ou quantitativa, tratando-se assim de uma tarefa de regressão (MUELLER; MASSARON, 2019).

2.3 Aprendizado Profundo

Uma subárea do aprendizado de máquina é o aprendizado profundo, que também trabalha com grandes conjuntos de dados para aprender a realizar tarefas. A grande diferença é que, ao invés de métodos e modelos estatísticos, o aprendizado profundo faz uso de apenas uma técnica que simula o funcionamento do cérebro humano: as redes neurais (MUELLER; MASSARON, 2019).

Redes neurais podem possuir diversas camadas e arquiteturas, sendo que cada uma pode ser mais ou menos indicada para cada tipo de problema a ser tratado. Tais redes são compostas por estruturas chamadas neurônios, que simulam os neurônios biológicos e se ligam uns aos outros por meio de ligações com diferentes pesos que simulam as sinapses cerebrais (MUELLER; MASSARON, 2019).

A rede neural mais simples possível é o *perceptron*, que pode ser utilizada para tarefas de classificação binária ou regressão linear. O *perceptron* é composto por apenas uma camada de um único neurônio. Ele recebe uma entrada X de tamanho m que é multiplicada por um vetor de pesos W também de tamanho m . Em seguida é aplicada sobre o produto escalar $X \cdot W$ uma função de ativação a fim de produzir um único resultado de saída (BLOCK, 1962). O processo de treinamento de uma rede neural consiste justamente em realizar, com base no conjunto de dados de entrada, a parametrização do vetor de pesos W das ligações entre os componentes da entrada X e o neurônio (MUELLER; MASSARON, 2019).

A partir do *perceptron*, novas arquiteturas de redes neurais podem ser formuladas adicionando mais camadas, mais neurônios, funções de retro-propagação de erro dentre outras estratégias que permitem que as redes neurais resolvam problemas mais complexos.

2.4 Transferência de aprendizado

A transferência de aprendizado é uma técnica utilizada em aprendizado de máquina na qual um modelo inicialmente treinado para executar uma determinada tarefa – tarefa de origem – é então retreinado ou ajustado para executar uma nova tarefa – tarefa alvo – similar à tarefa de origem (TORREY; SHAVLIK, 2010).

Dessa forma é possível reutilizar um modelo previamente desenvolvido. Outros benefícios da transferência de aprendizado são a melhora da performance – devido ao conhecimento previamente adquirido – e a eficiência dos dados, pois torna possível realizar o treinamento para a tarefa alvo com um conjunto de dados de tamanho reduzido. Sendo assim, a transferência de aprendizado é comumente utilizada quando se deseja realizar ta-

refas com domínio semelhante, ou também quando se possui poucos dados para se realizar o treinamento (TORREY; SHAVLIK, 2010).

2.5 Redes Neurais Convolucionais

Desde o ano de 2012, quando a arquitetura *AlexNet* venceu a competição *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)*, as Redes Neurais Convolucionais (RNC) vêm ganhando popularidade na execução de tarefas relacionadas a processamento de imagens, vídeos e até mesmo voz (VARGAS; PAES; VASCONCELOS, 2016). Esse tipo de rede neural é uma variação das redes de *perceptron* multicamada, e se baseia no princípio biológico da percepção visual dos seres humanos a fim de minimizar o processamento dos dados para obter o resultado esperado (MUELLER; MASSARON, 2019). Compostas por camadas convolucionais, de agrupamento, achatamento e conectada. Podem apresentar outros tipos de camadas além desses tipos básicos, ou ainda não apresentar alguma das camadas mencionadas acima dependendo do objetivo da rede (MUELLER; MASSARON, 2019). As próximas subseções detalham o papel de cada um desses elementos.

2.5.1 Camadas convolucionais

Na arquitetura de uma RNC, ilustrada na Figura 3, ao receber uma imagem como entrada são utilizadas as ditas camadas convolucionais. Tais camadas aplicam filtros (também conhecidos por *kernels*), representados por matrizes tridimensionais, sobre parte da imagem de entrada (RAWAT; WANG, 2017).

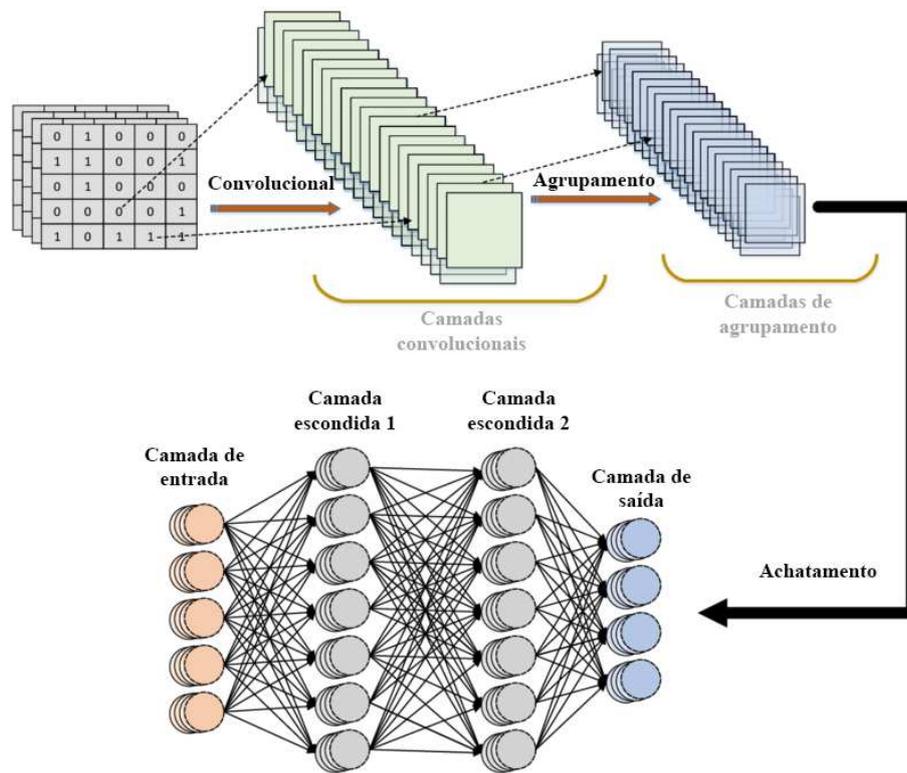


Figura 3 – Exemplo de estrutura básica de uma rede neural convolucional. Imagem adaptada de (SRINIVASAN et al., 2021)

A aplicação desses filtros se dá por um algoritmo de janela deslizante, que percorre a imagem em seus três canais de cores utilizando um passo de determinado tamanho (também conhecido como *stride*) realizando uma operação de produto escalar a fim de gerar uma nova matriz cujas dimensões podem ser menores ou iguais à anterior de acordo com o tamanho do filtro e do *stride*. As matrizes resultantes dessa operação entre os canais de cor e os filtros são chamadas de mapas de características. Esses são importantes parâmetros que devem ser determinados durante o projeto da rede. Esse é o processo chamado convolução, que é ilustrado na Figura 4 e é o mais importante processo das RNCs (GÉRON, 2019).

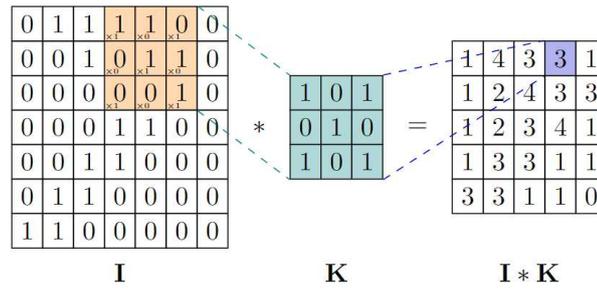


Figura 4 – Exemplo de convolução entre duas matrizes bidimensionais. O filtro K percorre a matriz I com um passo de uma unidade calculando os produtos escalares entre a respectiva região da matriz I e o filtro K, gerando assim uma nova matriz I*K que é o mapa de características, resultado da convolução entre I e K. Fonte: Autor.

2.5.2 Camadas de agrupamento

Após cada convolução é executada outra etapa importante do processamento em RNCs que é o agrupamento. Para esta tarefa existem as camadas de agrupamento (do inglês, *pooling*) cuja função é diminuir o tamanho do mapa de características visando agilidade e invariância espacial (RAWAT; WANG, 2017). Essa etapa consiste na utilização de uma janela deslizante a qual é aplicada uma função que seleciona um único valor para representá-la, conforme ilustrado na Figura 5. O uso de funções de valor máximo, mínimo e médio é bastante comum nesta etapa (GÉRON, 2019).

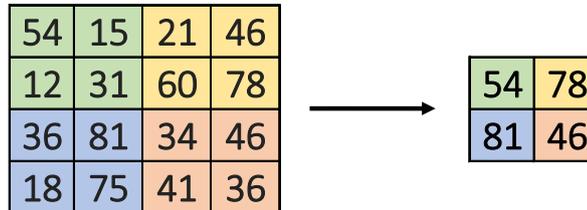


Figura 5 – Exemplo de agrupamento utilizando uma função de *max-pooling*. O mapa de características é percorrido por uma janela deslizante que seleciona apenas o maior valor da região para compôr um mapa de caractrísticas menor. Fonte: Autor.

Devido ao fato de diminuir o tamanho da imagem de entrada, essa etapa do processo – composta por camadas de convolução seguidas de camadas de agrupamento – também é conhecida como caminho de contração (GÉRON, 2019).

2.5.3 Camadas de achatamento

Após as etapas de convolução é executada uma camada de achatamento (do inglês, *flatten layer*), na qual os mapas de características, que são matrizes, são transformados em um vetor que será usado para alimentar uma rede de neurônios multicamada (GÉRON, 2019).

2.5.4 Camada densa ou conectada

A camada que comporta a rede de neurônios é chamada camada completamente conectada (ou densa) e tem por objetivo traçar a decisão a partir dos mapas de características. Esse tipo de camada é muito utilizado em tarefas de classificação (RAWAT; WANG, 2017).

2.6 U-Net

Proposta em 2015 por (RONNEBERGER; FISCHER; BROX, 2015), as redes U-Net vêm sendo aplicadas em diversos trabalhos referentes à segmentação de imagens médicas, uma vez que as redes de classificação, como as RNCs tradicionais, não conseguem trazer informações contextuais a nível de pixel, o que é essencial em tarefas de segmentação, especialmente ao se trabalhar com imagens médicas (SIDDIQUE et al., 2021).

As redes U-Net diferem-se das RNCs de classificação principalmente devido à adição de uma etapa de decodificação. Dessa forma, após passar pelo caminho de contração o vetor de saída é novamente transformado em uma matriz e esta é ampliada por camadas de aumento até chegar ao tamanho original da entrada, sendo a imagem resultante uma máscara que representa a região de interesse segmentada. Esse processo também é conhecido como caminho de expansão. A combinação do caminho de contração com o caminho de expansão resulta em uma rede quase simétrica, cujo formato lembra a letra U, conforme ilustrado na Figura 6, dando assim o nome a esse tipo de rede (RONNEBERGER; FISCHER; BROX, 2015).

Outra interessante característica desse tipo de rede é a ausência de uma camada conectada, sendo que esta é substituída por camadas convolucionais que processam os mapas de características com mais eficiência que uma camada conectada, conforme descrito em (LONG; SHELHAMER; DARRELL, 2015). A ausência da camada conectada, dá o nome à arquitetura das Redes Neurais Completamente Convolucionais (RNCC).

Assim, a arquitetura básica de uma U-Net, conforme foi originalmente proposta, consiste em um caminho de contração e um caminho de expansão. Durante a contração são aplicadas repetidamente duas convoluções com um filtro 3×3 seguidas de uma função de ativação *ReLU* e um *max-pooling* de tamanho 2×2 , diminuindo em duas vezes e aumentando em duas vezes, respectivamente, o tamanho e o número dos mapas de características da saída da camada anterior (SIDDIQUE et al., 2021). Na expansão é feito o processo inverso: a cada camada é feito o aumento de resolução do mapa de características com uma convolução 2×2 que diminui pela metade os canais de características e dobra as dimensões da imagem (*up-convolution*), uma concatenação com o mapa de características correspondente da etapa de contração e duas convoluções 3×3 seguidas por uma função de ativação *ReLU*. Por fim é feita uma convolução 1×1 para mapear vetores de características para as classes desejadas. A rede possui um total de 23 convoluções (RONNEBERGER; FISCHER; BROX, 2015).

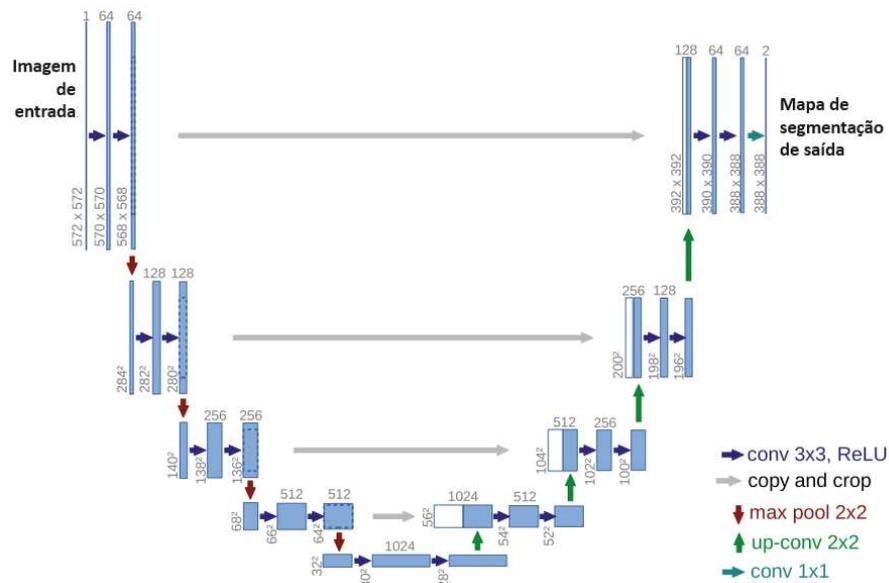


Figura 6 – Estrutura básica da rede U-Net. Fonte: (RONNEBERGER; FISCHER; BROX, 2015)

2.7 Métricas de validação

Esta seção apresenta e explica as principais métricas utilizadas para a validação do método proposto neste trabalho. Para isso, consideremos a seguinte nomenclatura:

- VP - Verdadeiros positivos
- VN - Verdadeiros negativos
- FP - Falsos positivos
- FN - Falsos negativos

2.7.1 Acurácia

A acurácia é a relação entre a quantidade de elementos classificados corretamente e a quantidade total de elementos analisados. Muitas vezes a acurácia não é a medida mais confiável para se avaliar modelos de aprendizado de máquina, pois é um desafio encontrar (e até mesmo construir) conjuntos de dados balanceados, ou seja, com aproximadamente a mesma quantidade de elementos de cada classe (UDDIN, 2019). Dessa forma, se um conjunto de dados para classificação binária possui um número muito mais elevado de elementos que não pertencem a uma classe $C1$ do que elementos que pertencem a $C1$, o modelo pode classificar corretamente os elementos que não pertencem a $C1$ e incorretamente uma boa parte dos elementos de $C1$ e ainda assim apresentar uma boa acurácia, visto que os elementos que não pertencem a $C1$ compõem a maior parte do conjunto de dados. A acurácia é calculada segundo a Equação 1:

$$acuracia = \frac{VP + VN}{VP + FN + FP + VN}. \quad (1)$$

2.7.2 Especificidade

A especificidade é a razão entre os elementos negativos que foram classificados corretamente, ou seja, os verdadeiros negativos VN; e os elementos que são de fato negativos, ou seja, a soma entre os verdadeiros negativos VN e falsos positivos FP. Novamente, se o número de elementos negativos do conjunto de dados for muito superior ao número de elementos positivos, a especificidade também pode ser uma métrica enviesada. A especificidade é calculada segundo a Equação 2.

$$especificidade = \frac{VN}{VN + FP}. \quad (2)$$

2.7.3 Precisão

A precisão é a razão entre o número de elementos classificados corretamente como positivos e o número de elementos que foram classificados como positivos (??), conforme mostra a Equação 3.

$$precisao = \frac{VP}{VP + FP}. \quad (3)$$

2.7.4 Sensibilidade

A sensibilidade (também conhecida como *recall*) é a razão entre o número de elementos positivos classificados corretamente e o número total de elementos que são de fato positivos. A sensibilidade é calculada segundo a Equação 4.

$$sensibilidade = \frac{VP}{VP + FN}. \quad (4)$$

2.7.5 F1-Score

O *f1-score* (também conhecida como *Dice Coefficient*) é a média harmônica entre precisão e a sensibilidade. Calculada conforme a Equação 5, busca representar a qualidade geral do modelo e é a métrica mais confiável quando se busca obter um modelo que minimize tanto os falsos positivos como os falsos negativos (??).

$$f1-score = \frac{2 * precisao * sensibilidade}{precisao + sensibilidade}. \quad (5)$$

2.7.6 Intersecção sobre União

A Intersecção sobre União (*Intersection over Union – IoU*, também conhecida como *Jaccard Index*) é uma medida para calcular a similaridade entre conjuntos (RAHMAN; WANG, 2016) e é dada pela Equação 6.

$$IoU = \frac{VP}{VP + FP + FN}. \quad (6)$$

2.8 Trabalhos Relacionados

Esta seção apresenta os principais trabalhos relacionados a esta pesquisa e que mostram técnicas de segmentação em imagens histológicas.

A literatura apresenta poucos trabalhos focados em segmentação de canais ósseos em imagens de lâmina inteira. Em (LIU et al., 1999) a segmentação foi realizada em imagens micro-radiográficas usando conhecimento em imagens ósseas e algoritmos de agrupamento. As imagens de lâmina inteira diferem-se das imagens micro-radiográficas especialmente em tamanho e resolução já que podem assumir ordens de giga-pixel e apresentar cores enquanto as imagens utilizadas por (LIU et al., 1999) possuíam dimensões de 512×576 pixels e se apresentavam em escala de cinza. A alta resolução das imagens de lâmina inteira é interessante na análise histológica devido à riqueza de detalhes que proporciona ao especialista. Porém, trabalhar com tais imagens é uma tarefa desafiadora devido ao seu tamanho, complexidade e quantidade das estruturas em relação às imagens micro-radiográficas (GONDIM et al., 2021).

Um método para segmentação automática de canais ósseos em imagens histológicas de tecido ósseo foi proposto por (GONDIM et al., 2021). Tal método processou as imagens de entrada, resumidamente, em quatro etapas: remoção de artefatos externos à matriz óssea, remoção de artefatos internos à matriz óssea, segmentação da rede de canais ósseos e remoção de falsos positivos. A cada etapa foram utilizadas diversas técnicas de processamento de imagens (tais como abertura e fechamento morfológico, binarização, deconvolução de cor, etc.) em sequência e até mesmo um algoritmo de aprendizado de máquina não supervisionado denominado *K-means*. Isso tornou o método bastante complexo e implica em uma grande quantidade de parâmetros cujos valores devem ser fixos e obtidos empiricamente, o que leva a questionar sobre sua robustez em relação a variações na imagem. Na Figura 7 é mostrado um exemplo de entrada e saída do método proposto por (GONDIM et al., 2021).

Para a avaliação do método, as regiões segmentadas por (GONDIM et al., 2021) foram comparadas com marcações feitas por um especialista. Na metodologia de (GONDIM et al., 2021), como as marcações manuais não seguiam a forma exata das estruturas, para comparar a segmentação do método foi considerado que se ao menos 20% dos pixels de uma marcação feita pelo especialista estivessem contidos na região segmentada pelo

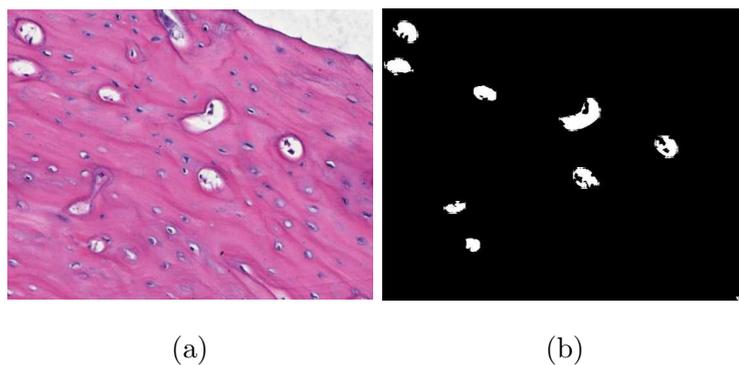


Figura 7 – Exemplo de entrada e saída do método proposto por (GONDIM et al., 2021). Em (a) imagem de entrada. Em (b) a saída do método para a entrada em (a). Fonte (GONDIM et al., 2021).

método, então a estrutura foi corretamente segmentada. Isso deixa uma lacuna em relação à forma da região segmentada pois tais formas não puderam ser validadas devido às marcações manuais, pois estas foram feitas colocando um círculo envolvendo cada canal, como mostrado na Figura 8.

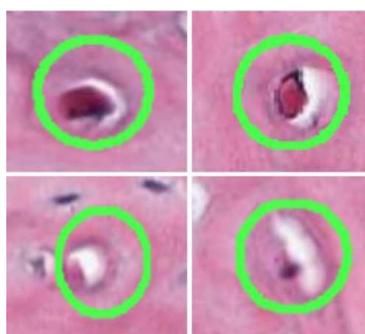


Figura 8 – Exemplo de marcações feitas pelo especialista. Fonte (GONDIM et al., 2021).

Outra abordagem foi testada por (JULIA et al., 2021), que, na intenção de segmentar canais ósseos em imagens histológicas geradas a partir do fêmur de ratos, modificou uma rede neural convolucional do tipo U-Net que originalmente havia sido feita para segmentação de anormalidades FLAIR em imagens de ressonância magnética cerebral. Entretanto os resultados da marcação automática feita pela RNC implementada não foram satisfatórios devido à falta de precisão das marcações manuais feitas nas imagens, o que é ilustrado na Figura 9. O modelo treinado apresentou sobreajuste e alcançou um *Dice Coefficient* de apenas 20%.

Como proposta de melhoria em seu trabalho, (JULIA et al., 2021) sugere o uso de um conjunto de imagens com marcações mais precisas, a fim de melhorar os resultados da RNC.

Em ambos os trabalhos observa-se que a qualidade do conjunto de imagens teve impacto negativo na avaliação dos resultados da segmentação. Acredita-se que um conjunto de imagens com marcações mais precisas e confiáveis em relação ao tamanho, forma e po-

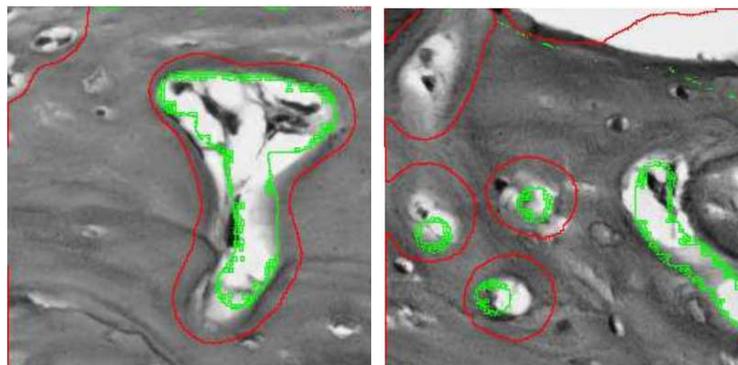


Figura 9 – Exemplos de saídas do método proposto. Em verde a saída da rede neural, em vermelho a marcação provida pelo especialista. Imagem adaptada de (JULIA et al., 2021).

sição das estruturas de interesse pode levar a um resultado mais preciso, e o uso de redes neurais convolucionais pode resultar em um método mais robusto quanto à invariância, apresentando assim mais flexibilidade ao tratar imagens com características diferentes.

Pensando nos bons resultados que as redes neurais convolucionais têm apresentado no processamento de imagens nos últimos anos, (SANTOS et al., 2021) utilizou redes neurais convolucionais para segmentar regiões tumorais em imagens de lâmina inteira de tecido oral, obtendo bons resultados, como cerca de 97,6% de acurácia, 98,4% de especificidade e 92,9% de sensibilidade. Foi utilizada uma RNCC baseada no modelo U-Net treinada com uma base de imagens, desenvolvida pelos autores, que continha imagens histológicas de tecidos da cavidade oral corados com HE.

Antes do treinamento foi feita uma etapa de pré-processamento para identificar as regiões de tecido, descartando áreas desinteressantes para o trabalho. As imagens eram então divididas em sub imagens e apresentadas à rede para o treinamento, que também contava com uma estratégia de aumento de dados. O resultado da rede era uma imagem em escala de cinza em que era calculada para cada pixel a probabilidade p de tal pixel pertencer ou não à região de interesse. Por fim eram marcados como positivos os pixels que apresentassem p acima de um limiar de 50%, valor este que foi escolhido empiricamente. Na Figura 10 se encontram exemplos dos principais elementos gráficos presentes no método proposto por (SANTOS et al., 2022).

O método também foi validado em outras bases de imagens presentes na literatura, e apresentou um $f1$ -score de 90% na base desenvolvida pelos autores e média geral de 83%. Além disso o método foi testado com diferentes tamanhos de sub imagens – o melhor resultado foi obtido para sub imagens de tamanho 640×640 pixels.

Portanto observou-se um potencial no uso de RNCs para realizar segmentação de regiões de interesse em imagens histológicas de lâmina inteira coradas com HE. Porém para a segmentação de canais ósseos a ausência de um conjunto de imagens adequado e com boas marcações mostrou-se um problema, tanto para treinamento de RNCs como

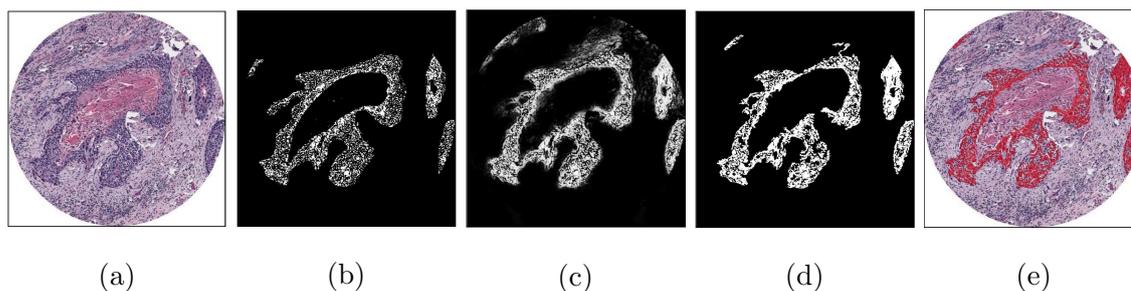


Figura 10 – Exemplos dos principais elementos gráficos presentes no método proposto por (SANTOS et al., 2022). Em (a) uma imagem de entrada. Em (b) um exemplo de marcação padrão ouro. Em (c) um exemplo da saída da rede, em que cada pixel é preenchido de acordo com a probabilidade de pertencer à região de interesse (quanto mais próximo da cor branca, maior a probabilidade). Em (d) a marcação final do método, com *limiar* de 50%. Em (e) a região com tumor representada em vermelho na imagem original. Imagem adaptada de (SANTOS et al., 2022).

para validação da metodologia.

Materiais e Métodos

Este capítulo descreve o conjunto de dados elaborado, a arquitetura da rede neural e o método de segmentação proposto. Na Figura 11 é apresentado um esquema geral da metodologia utilizada.

3.1 Conjunto de dados

As imagens utilizadas foram obtidas a partir do fêmur esquerdo de um rato *Rattus norvegicus* da linhagem *Wistar* saudável com 200 - 250g. O animal foi sacrificado e os fêmures foram removidos e fixados em formaldeído 10%, tamponado e desmineralizado em EDTA 4,13%. A diáfise (Figura 1), região intermediária do fêmur, foi incluída em parafina e obteve-se cortes histológicos seriados de cerca de 5 micrômetros de espessura corados em HE, que é o tipo de coloração mais amplamente utilizado na histologia devido à sua simplicidade e ao grande número de estruturas do tecido que permite visualizar (FELDMAN; WOLFE, 2014).

Todos os procedimentos para se obter este material foram executados de acordo com as normas do Colégio Brasileiro de Experimentação Animal (COBEA), com aprovação do Comitê de Ética no Uso de Animais da Universidade Federal de Uberlândia (CEUA-UFU) – Protocolo 060/09.

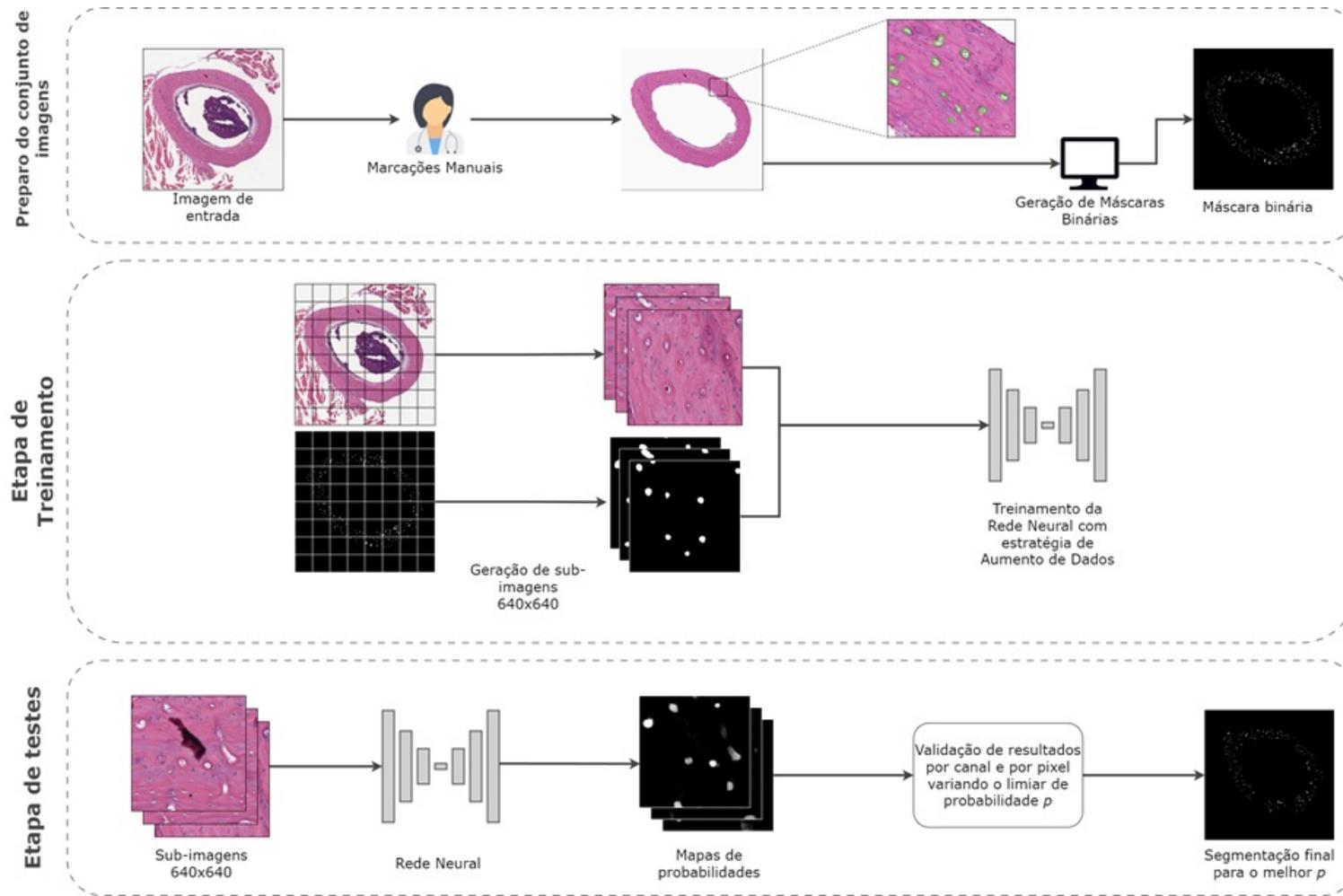
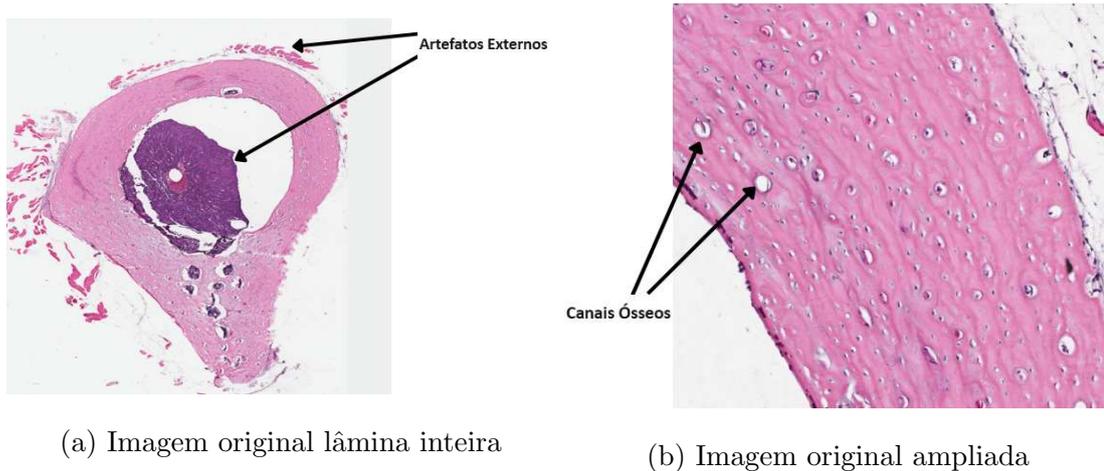


Figura 11 – Metodologia utilizada. Inicialmente as imagens foram marcadas manualmente. Máscaras binárias foram geradas destacando as estruturas de interesse. A rede neural foi então treinada utilizando sub imagens das imagens originais e suas máscaras; o treinamento contou com uma estratégia de aumento de dados. Mapas de probabilidades foram computados utilizando a rede treinada; diferentes valores de limiar de probabilidade p foram testados em dois tipos de análise: pixel a pixel e canal a canal; a segmentação final foi feita utilizando o valor de p que apresentou melhores resultados. Fonte: Autor.

Neste estudo foram analisados 65 cortes histológicos de fêmur esquerdo. Foi utilizado um ScanScope AT Turbo®Scanner (Leica Biosystems, Nussloch, Alemanha) para digitalizar as imagens com uma ampliação efetiva de $20\times$. As imagens foram exportadas em arquivos TIFF de alta resolução. Na Figura 12 é mostrado um exemplo de corte histológico para o fêmur irradiado. O tamanho da imagem na Figura 12(a) é de 7.940×9.051 pixels, a largura e altura do pixel são $0,502 \mu\text{m}$. Nessa resolução é possível identificar os canais ósseos (Figura 12(b)).



(a) Imagem original lâmina inteira

(b) Imagem original ampliada

Figura 12 – Exemplo de imagem utilizada no método proposto. Em (a) imagem da secção inteira destacando artefatos externos que não são de interesse na análise, em (b) imagem ampliada destacando os canais ósseos. Fonte: Autor.

Todas as imagens coletadas foram marcadas por um especialista em histologia. A marcação foi feita por meio do software Photoshop®, versão 2016, contornando manualmente os canais ósseos com a ferramenta ‘Laço’. Em seguida foram exportadas em formato JPEG. Nesta etapa também foi feita a remoção do fundo da imagem, removendo artefatos externos que não eram interessantes para a análise devido ao fato de que a estrutura de interesse não se apresenta em tais artefatos. Os cortes histológicos marcados manualmente foram posteriormente divididos em 2.037 sub imagens de dimensões 640×640 pixels conforme será descrito adiante no texto. Na Figura 13(a) é apresentado um exemplo de imagem marcada, e na Figura 13(b) uma região ampliada da mesma imagem evidenciando as marcações.

As imagens de lâmina inteira bem como as sub imagens utilizadas neste trabalho estão disponíveis publicamente no endereço: <https://www.kaggle.com/datasets/igorgonribsilva/histological-bone-canals-wsi>.

3.2 Arquitetura da Rede Neural

Assim como em (SANTOS et al., 2022), o objetivo deste trabalho é realizar segmentação de imagens de lâmina inteira coradas com HE. Portanto, optou-se por utilizar a mesma

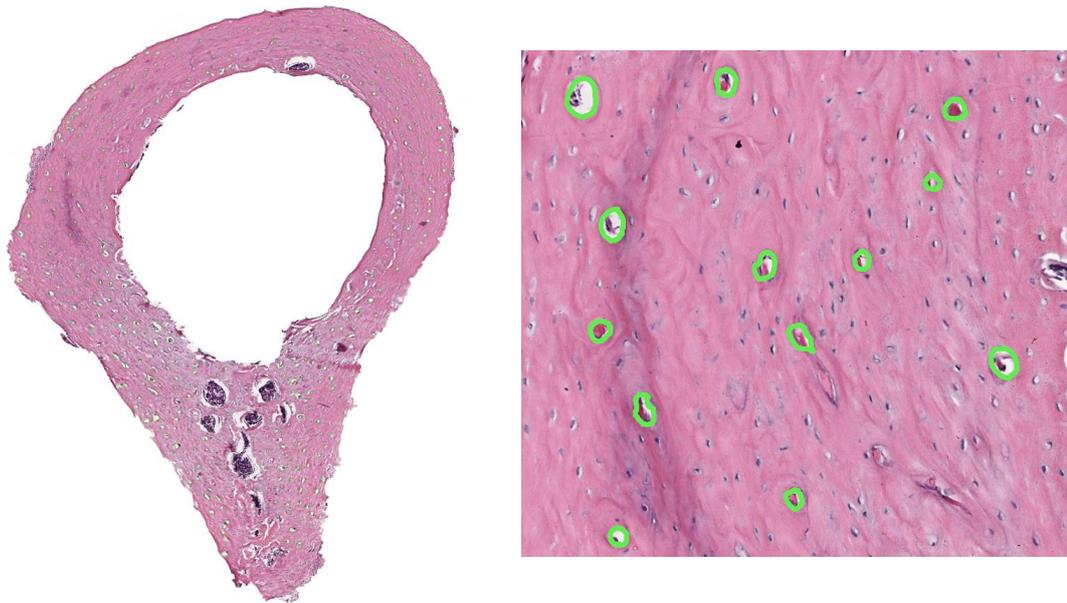


Figura 13 – Imagem marcada por especialista. À esquerda a imagem inteira. À direita a imagem ampliada destacando os canais ósseos, contornados em verde. Fonte: Autor.

rede neural. Não foram feitas alterações na rede desenvolvida por (SANTOS et al., 2022), que por sua vez é baseada na rede U-Net originalmente proposta por (RONNEBERGER; FISCHER; BROX, 2015) com algumas modificações que estão destacadas na Figura 14 e serão descritas a seguir:

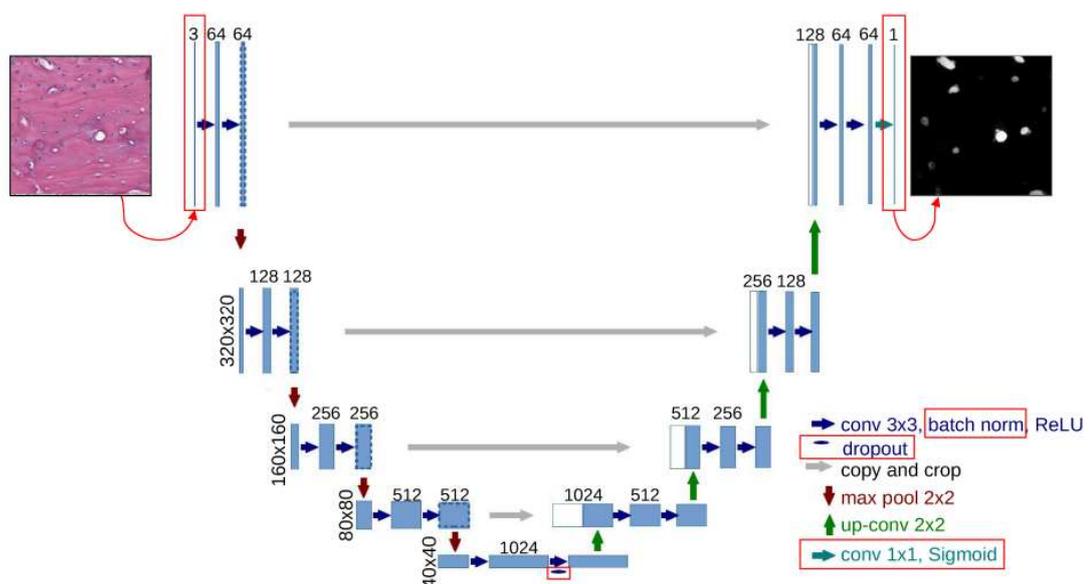


Figura 14 – Arquitetura da rede utilizada baseada na rede U-Net original com modificações destacadas em vermelho. Imagem adaptada de (SANTOS et al., 2022).

- Normalização em lote: A normalização em lote normaliza as entradas evitando valores muito baixos que possam gerar gradiente de fuga e fazer com que os pesos

não se atualizem (ou se atualizem muito lentamente) durante o treinamento. A técnica também evita valores muito altos que possam gerar explosão de gradiente fazendo com que os pesos assumam valores muito altos rapidamente prejudicando o treinamento (GÉRON, 2019). A normalização em lote foi aplicada após cada camada convolucional.

- *Dropout*: O *dropout* é uma técnica que consiste em desativar parte dos neurônios da camada conectada durante algumas etapas do treinamento a fim de evitar o sobreajuste, tornando a rede mais robusta e generalizável (GÉRON, 2019). Foi utilizada uma taxa de *dropout* de 0,5. Dessa forma a cada passo do treinamento metade dos neurônios eram desativados aleatoriamente para minimizar as chances de sobreajuste.
- Convolução 1x1 com função de ativação Sigmoid: Ao final do processamento, foi adicionada à última convolução (1x1) uma função de ativação Sigmoid a fim de que, ao invés de uma imagem binária, a saída da rede fosse uma imagem em escala de cinza onde quanto mais próxima de branco a cor de um pixel maior a probabilidade de o mesmo pertencer à região de interesse.

3.3 Método

A fim de reaproveitar a rede neural proposta por (SANTOS et al., 2022), foi necessário transformar as imagens marcadas manualmente em imagens binárias que destacassem as regiões de interesse. Na Figura 15 é mostrada uma marcação manual e a respectiva máscara binária gerada.

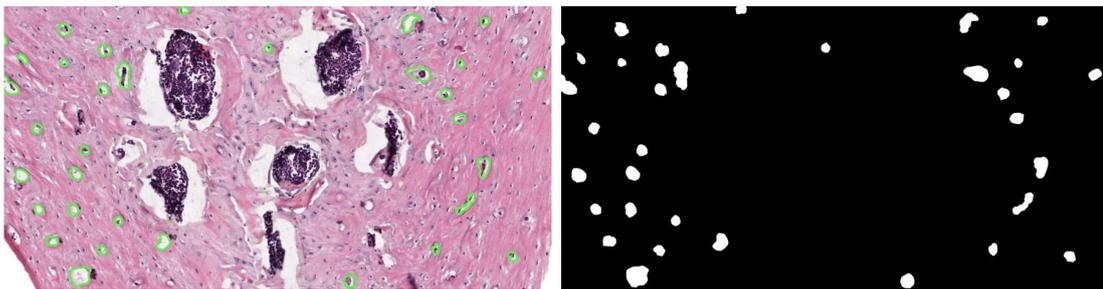


Figura 15 – Região de imagem marcada manualmente (a) e sua respectiva máscara binária (b). Fonte: Autor.

O fundo das imagens originais (sem marcações) foi removido utilizando do processo descrito por (SANTOS et al., 2022): diminuindo-se a escala das imagens por um fator de 32x, destacando-se a região com tecido por meio de um filtro de cor e mapeando-se a região selecionada na imagem em tamanho real.

Ainda como foi feito por (SANTOS et al., 2022), foram geradas sub imagens a partir das imagens originais, desconsiderando-se as regiões que não contêm tecido. O tamanho

de sub imagens utilizado foi 640×640 pixels, visto que foi o tamanho que apresentou melhores resultados em (SANTOS et al., 2022). As sub imagens que continham apenas fundo foram descartadas.

O treinamento ainda aplicou uma estratégia de aumento de dados utilizada por (SANTOS et al., 2023), em que a cada época é aplicada uma transformação selecionada aleatoriamente. Para este aumento de dados, as seguintes operações foram consideradas: Inversão Horizontal, Inversão Vertical, Rotação, Tranposição, Transformação elástica, Distorção de grade e Distorção ótica. Na Figura 16 são mostrados exemplos de cada uma dessas operações em comparação com um exemplar de sub-imagem original.

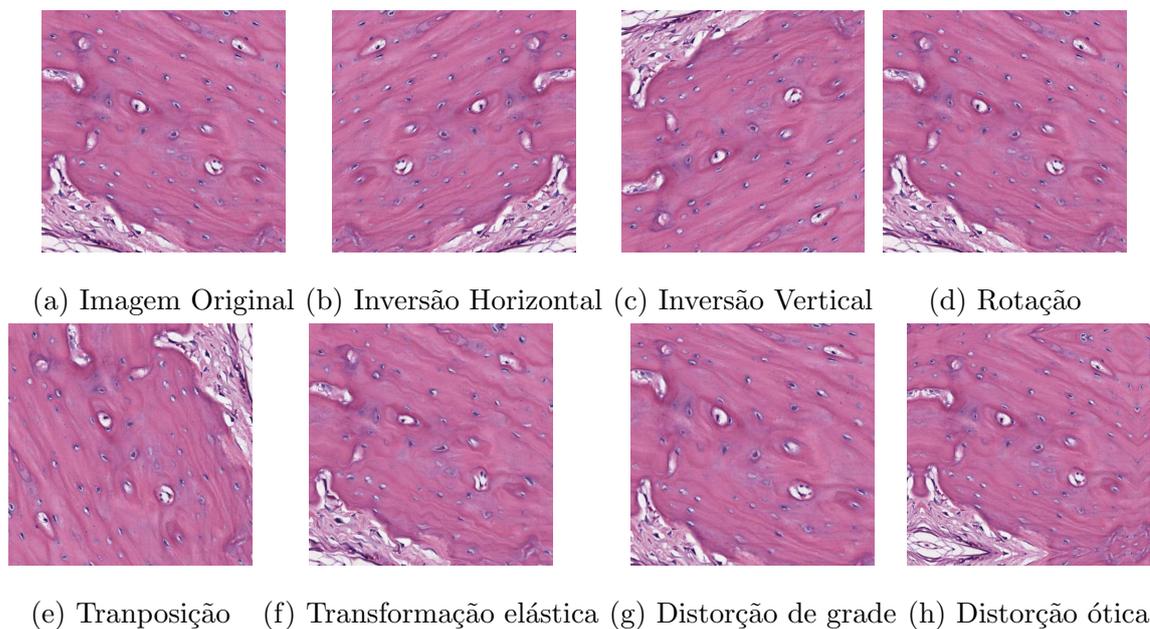


Figura 16 – Transformações utilizadas no aumento de dados. Fonte: Autor.

O conjunto de dados de treinamento foi composto por 80% das sub imagens, destas, 20% foram usados para validação da rede durante o treinamento. Os 20% restantes compuseram o conjunto de teste.

Após o treinamento foi realizada a etapa de inferência, na qual foram apresentadas à rede imagens do conjunto de testes, ou seja, que não foram utilizadas na etapa de treinamento, para que a segmentação dos canais fosse feita e as métricas calculadas.

Nesta etapa foi calculada para cada pixel uma probabilidade de que o mesmo pertença a um canal. Caso a probabilidade seja maior que uma probabilidade limiar p , entende-se que aquele pixel faz parte da região de interesse. Para determinar qual o valor a ser usado como probabilidade limiar foram testados para este parâmetro valores de 5% a 95% com um passo de 5%. Para cada imagem do conjunto de testes foram calculados acurácia, precisão, $f1$ -score, sensibilidade e especificidade. A média das métricas citadas acima foi calculada para cada valor de p testado a fim de identificar empiricamente qual o melhor valor de p a ser utilizado para a segmentação final.

Para o cálculo dessas métricas foram considerados: VP: pixel pertencente ao canal ósseo e detectado como pertencente ao canal ósseo; VN: pixel não pertencente ao canal ósseo e detectado como não pertencente ao canal ósseo; FP: pixel não pertencente ao canal ósseo e detectado como pertencente ao canal ósseo; FN: pixel pertencente ao canal ósseo e detectado como não pertencente ao canal ósseo.

Também foi feita uma análise das medidas precisão, *f1-score*, sensibilidade e intersecção sobre união para cada canal segmentado utilizando o método proposto por (GONDIM et al., 2021). Esse resultado serviu como comparação do método proposto com um método existente na literatura.

A fim de comparação, foram realizados dois experimentos seguindo a metodologia descrita acima: o primeiro foi o treinamento executado a partir de uma inicialização aleatória dos pesos da rede neural, a partir de agora o texto irá se referir a este treinamento como Treinamento a partir do zero (TZ); o segundo foi um treinamento executado a partir do modelo treinado e disponibilizado publicamente por (SANTOS et al., 2022), a partir de agora o texto irá se referir a este treinamento como Treinamento por Transferência de Aprendizado (TTA).

Ambos os experimentos foram realizados em um computador com processador Intel(R) Core(TM) i7-12700 2.10 GHz, 12^a geração; 16Gb de RAM; unidade de processamento gráfico NVIDIA GeForce RTX 3060; e sistema operacional Windows 11 Pro versão 22H2. Para o desenvolvimento do código foi criado um ambiente virtual utilizando o gerenciador de pacotes Anaconda e a linguagem de programação Python na versão 3.10.11.

Resultados e Discussão

Este capítulo apresenta os resultados obtidos nos experimentos descritos no capítulo anterior.

4.1 Treinamento da Rede

Para a fase de treinamento, embora os resultados obtidos com o TTA tenham sido ligeiramente inferiores que os obtidos em TZ, os resultados foram semelhantes, e não houve mudança significativa na curva de aprendizado da rede. Para ambos os treinamentos foram obtidos sobre o conjunto de validação acurácia acima de 97% e erro em torno de 12%. Na Figura 17 são mostradas as curvas de perda e acurácia para ambos os conjuntos de dados ao longo de ambos os treinamentos.

A acurácia apresentou valores elevados principalmente devido ao desbalanceamento do conjunto de dados, pois as imagens possuem muito mais pixels negativos do que positivos, visto que os canais ósseos são estruturas pequenas que ocupam pouco espaço nas imagens quando comparamos à matriz óssea, por exemplo.

4.2 Análise por pixel

Em seguida foi executada a inferência sobre as imagens do conjunto de testes. Nesta etapa a saída da rede neural é uma imagem em escala de cinza onde cada pixel possui um valor entre 0 e 255 e quanto maior a probabilidade de um pixel pertencer à região de interesse maior será o seu valor. Portanto, visualmente, quanto mais próximo da cor branca for o pixel na imagem de saída maior a probabilidade de o mesmo pertencer à região de interesse, conforme ilustrado na Figura 18.

Dessa forma, um pixel é considerado um positivo se a probabilidade de o mesmo pertencer à região de interesse for maior que um limiar p .

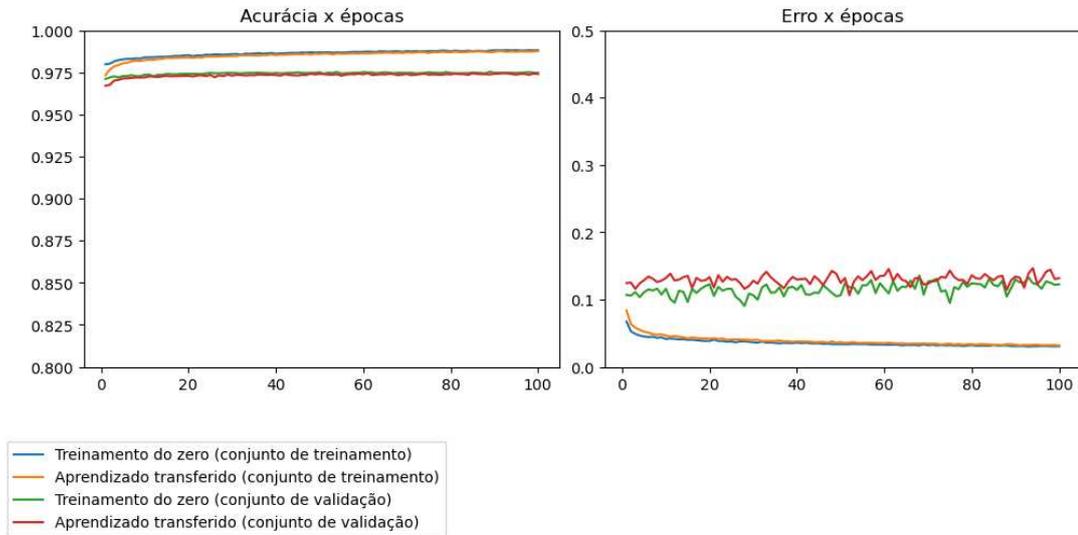


Figura 17 – Curvas de acurácia e perda ao longo de ambos os treinamentos. À esquerda a acurácia das previsões da rede aplicadas ao conjunto de treinamento (em azul para o TZ e em laranja para o TTA) e ao conjunto de validação (em verde para o TZ e em vermelho para o TTA). À direita a perda das previsões da rede aplicadas ao conjunto de treinamento (em azul para o TZ e em laranja para o TTA) e ao conjunto de validação (em verde para o TZ e em vermelho para o TTA). Fonte: Autor.

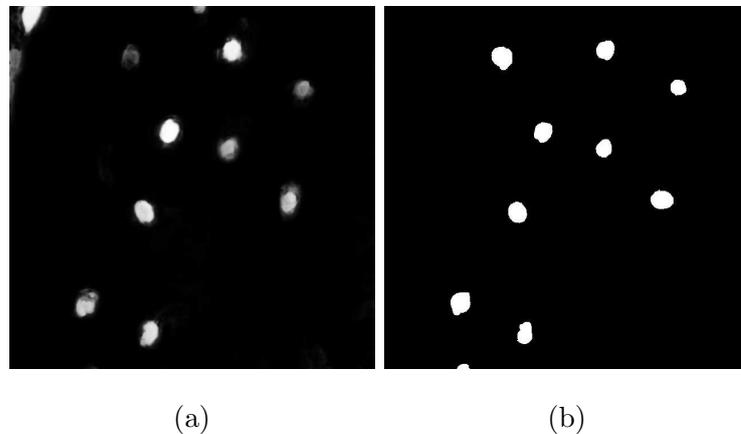


Figura 18 – Em (a) um exemplo de uma saída da RNC. Em (b) a respectiva máscara manualmente marcada pelo especialista. Fonte: Autor.

$$Y(x) = \begin{cases} 0, & \text{se } \frac{x}{255} < p \\ 1, & \text{caso contrário} \end{cases} \quad (7)$$

Para validar o melhor valor para tal limiar foram testados vários valores de p , variando-o de 0,05 até 0,95 com um passo de 0,05. Para cada valor testado foram calculados os valores médios de acurácia, precisão, $f1$ -score, sensibilidade e especificidade. A Tabela 1 e o gráfico na Figura 19 mostram os valores médios obtidos para cada valor de p .

Após o teste foi observado que o TTA obteve resultados melhores que o TZ, alcançando

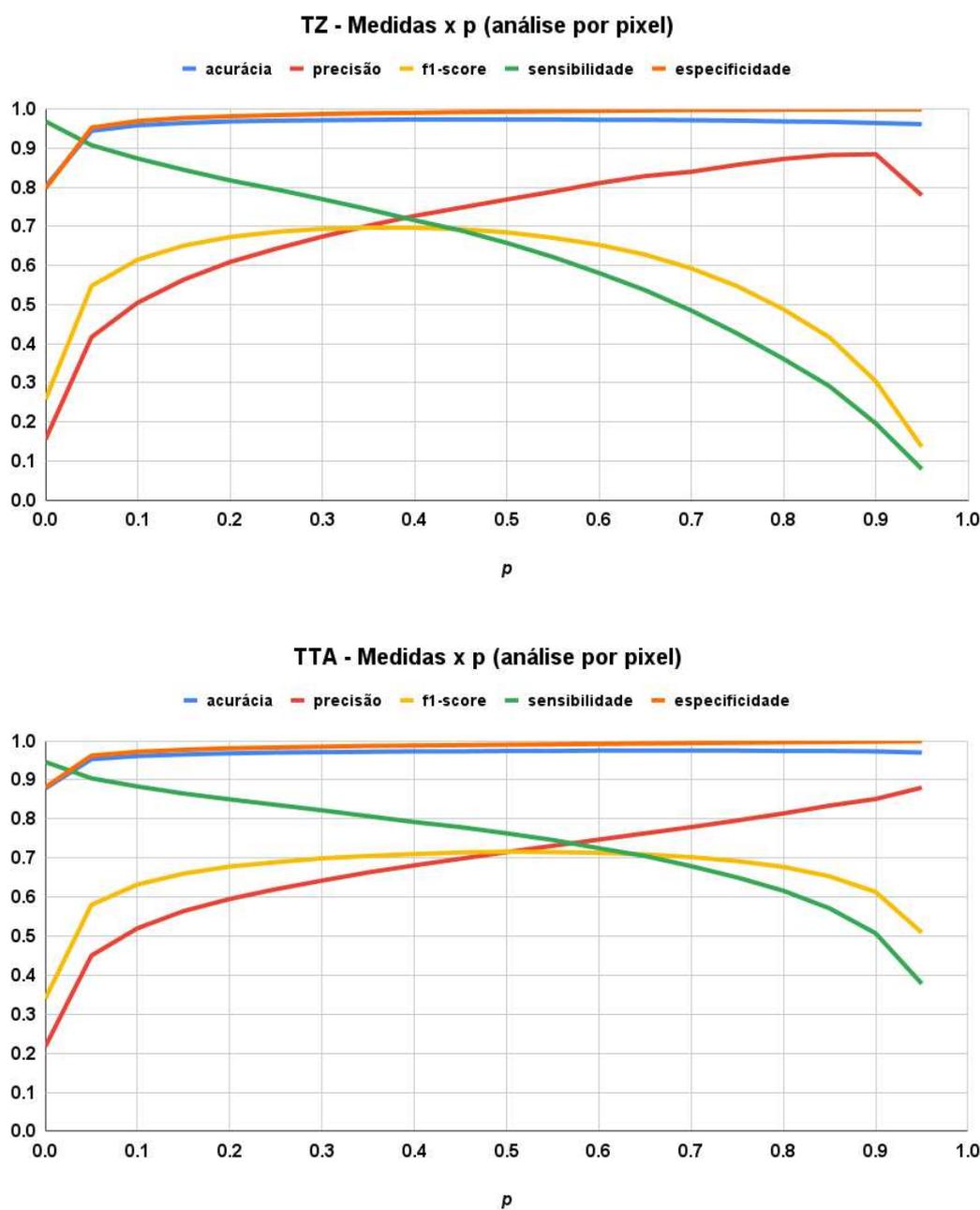


Figura 19 – Acurácia, $f1$ -score, precisão, sensibilidade e especificidade em função do th - $reshold$ p na análise por pixel de ambos os treinamentos. Fonte: Autor.

p	Treinamento do Zero					Transferência de aprendizado				
	Acur.	Prec.	F1	Sensib.	Espec.	Acur.	Prec.	F1	Sensib.	Espec.
0,05	0,944	0,416	0,548	0,907	0,953	0,953	0,450	0,580	0,904	0,962
0,10	0,958	0,505	0,615	0,873	0,969	0,961	0,520	0,632	0,883	0,972
0,15	0,965	0,563	0,651	0,844	0,977	0,965	0,564	0,660	0,865	0,977
0,20	0,968	0,608	0,673	0,817	0,982	0,968	0,595	0,678	0,850	0,981
0,25	0,970	0,642	0,685	0,794	0,985	0,970	0,620	0,689	0,836	0,983
0,30	0,971	0,674	0,693	0,769	0,987	0,971	0,642	0,699	0,822	0,985
0,35	0,972	0,701	0,697	0,743	0,989	0,972	0,663	0,705	0,807	0,987
0,40	0,973	0,726	0,696	0,716	0,991	0,973	0,681	0,710	0,792	0,988
0,45	0,973	0,747	0,693	0,689	0,992	0,973	0,698	0,714	0,779	0,989
0,50	0,973	0,768	0,684	0,657	0,993	0,974	0,715	0,716	0,763	0,990
0,55	0,973	0,789	0,671	0,621	0,995	0,974	0,731	0,715	0,746	0,991
0,60	0,973	0,810	0,652	0,580	0,996	0,975	0,747	0,713	0,725	0,992
0,65	0,972	0,829	0,628	0,537	0,996	0,975	0,763	0,709	0,705	0,993
0,70	0,971	0,840	0,592	0,484	0,997	0,975	0,779	0,702	0,679	0,994
0,75	0,970	0,858	0,547	0,426	0,998	0,975	0,796	0,692	0,650	0,995
0,80	0,969	0,873	0,487	0,360	0,998	0,974	0,814	0,677	0,616	0,996
0,85	0,967	0,882	0,415	0,290	0,999	0,974	0,834	0,653	0,571	0,997
0,90	0,965	0,884	0,303	0,195	0,999	0,973	0,851	0,613	0,507	0,998
0,95	0,962	0,780	0,135	0,078	0,999	0,970	0,880	0,509	0,378	0,999

Tabela 1 – Médias de acurácia, precisão, $f1$ -score, sensibilidade e especificidade para cada $threshold$ p testado na análise por pixel.

71,6% para $p=0,5$ de $f1$ -score, logo, a melhor relação entre precisão e sensibilidade. A especificidade e acurácia se mantiveram elevadas durante todo o teste devido ao alto volume de pixels negativos corretamente classificados.

Utilizando o resultado acima foi feita a segmentação das imagens do conjunto de testes utilizando o modelo treinado a partir do zero com $p=0,35$. Na Figura 20 se encontra a comparação entre a máscara gerada pela rede e a máscara gerada a partir das marcações do especialista, e na Figura 21 é mostrada uma imagem da lâmina inteira marcada a partir da máscara gerada pela rede e algumas regiões ampliadas para detalhar os resultados, em que ainda é possível observar alguns falsos negativos.

4.3 Análise por canal

Nesta análise, para cada componente conectado da máscara de referência foi extraída uma sub-imagem de acordo com as coordenadas x_i, x_f, y_i, y_f do componente, em que x_i e y_i representam o menor valor de coordenada do pixel em x e y, enquanto x_f e y_f representam o maior valor de coordenada do pixel em x e y. Extraía-se também uma sub-imagem da respectiva região delimitada por x_i, x_f, y_i, y_f na imagem segmentada pelo método para verificar se havia ou não um canal naquela região. Em seguida foi calculada a intersecção entre as duas sub imagens. Foram considerados como verdadeiros positivos os componentes cuja intersecção coincidiram ao menos 70% com o canal da

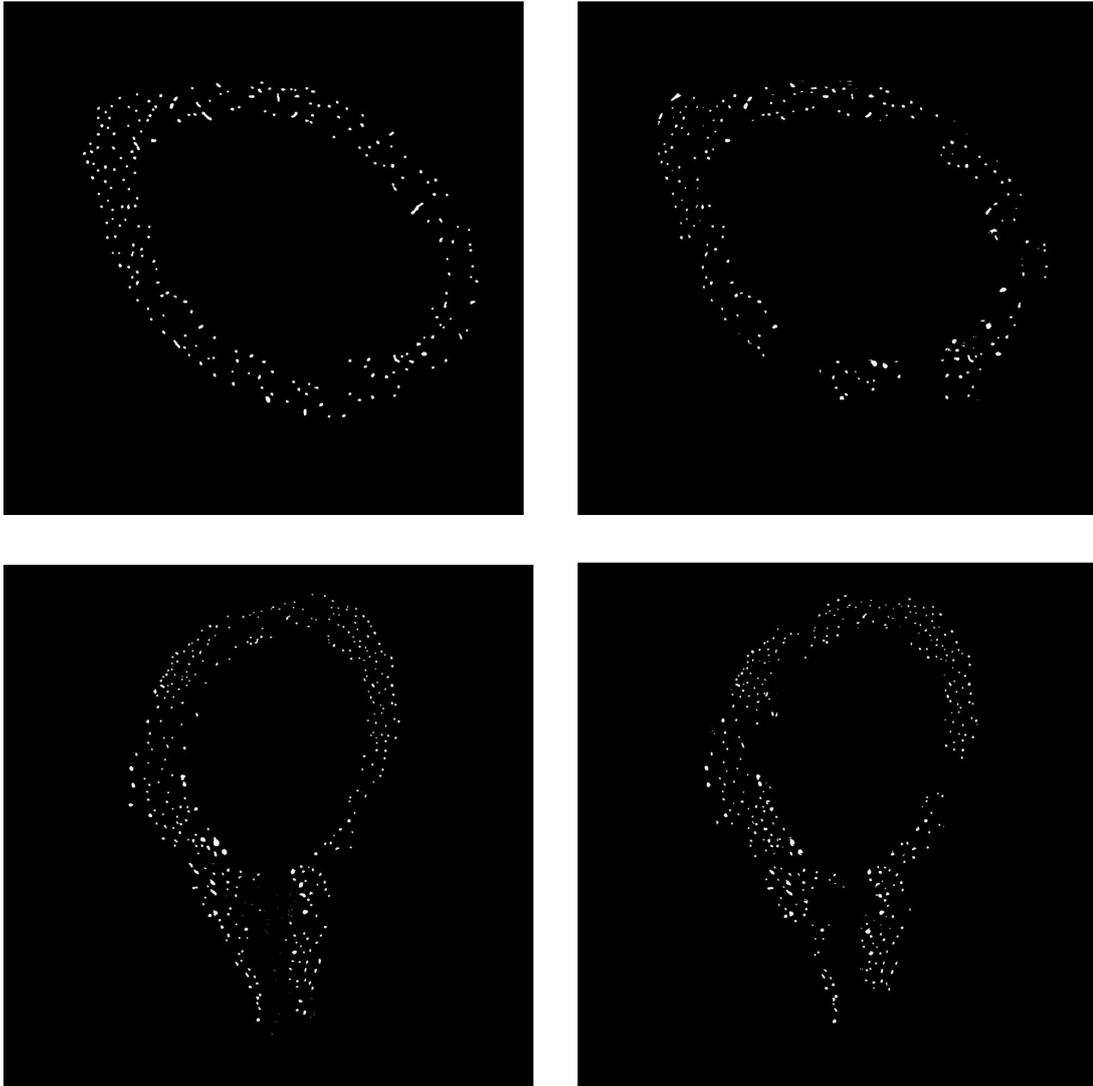


Figura 20 – Comparação entre marcação manual feita por especialista e saída do método. Na primeira coluna exemplos de marcações manuais, na segunda coluna as máscaras geradas pelo método treinado a partir do zero e com um limiar $p=0,35$. Imagem em resolução original disponível neste link. Fonte: Autor.

máscara de referência. O valor de 70% foi escolhido como um valor arbitrário superior ao valor utilizado como critério de aceite em (CARDOSO; CORTE-REAL, 2005), que foi de 50%. Na Figura 22 são mostrados exemplos de canais na máscara de referência, na segmentação da rede e a intersecção entre os dois canais.

Também foram extraídas da imagem segmentada sub imagens representando os componentes conectados que não estavam presentes na imagem de referência, identificando, portanto, os falsos positivos. Nessa análise não é possível contabilizar verdadeiros negativos.

Para essa nova análise foram testados novamente os valores do limiar de probabilidade p . Por não termos o número de verdadeiros negativos as métricas utilizadas foram: Precisão, $f1$ -score, sensibilidade e Intersecção sobre União. O melhor resultado para ambos

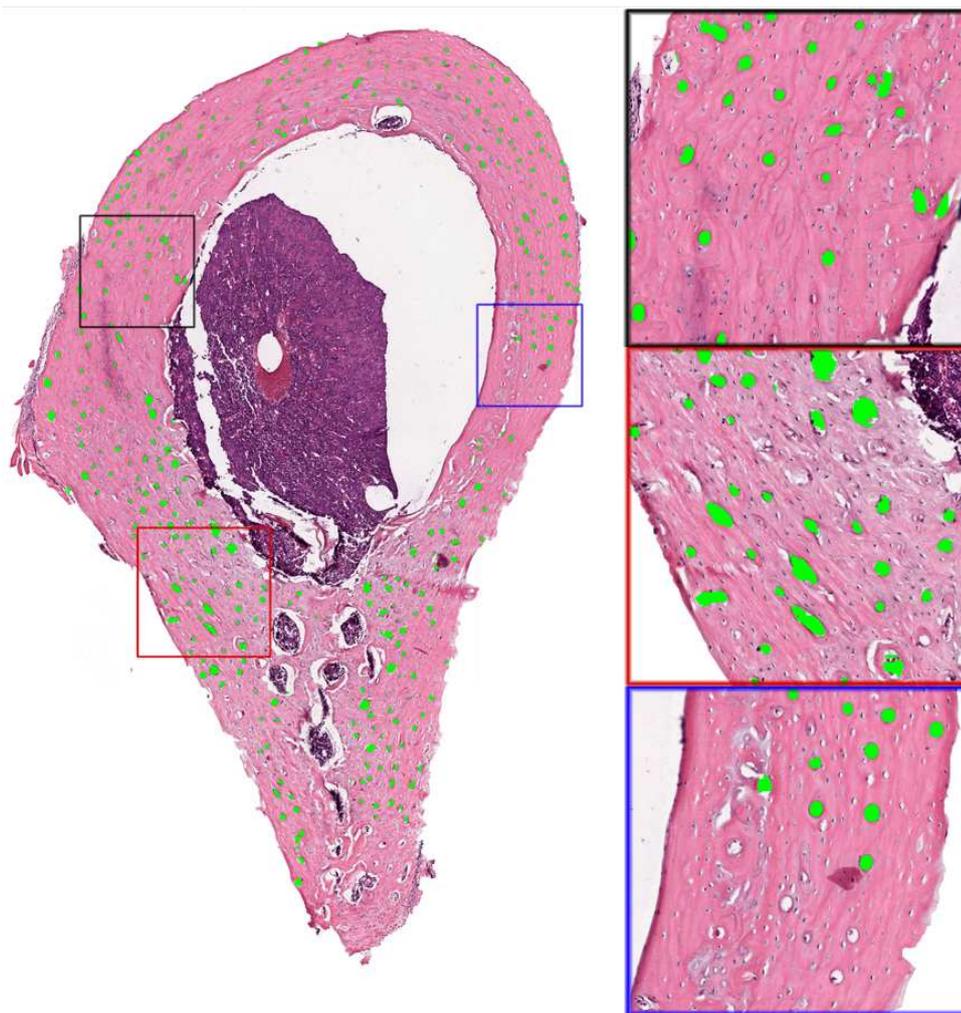


Figura 21 – Imagem de lâmina inteira com as marcações feitas pela rede em verde e com três regiões ampliadas. Observa-se, principalmente na região delimitada pela cor azul, a presença de alguns falsos negativos. Fonte: Autor.

os treinamentos foi obtido para $p = 0, 10$. A Tabela 2 mostra o resultados da análise para cada valor de p em ambos os treinamentos.

A análise descrita acima foi realizada sobre as imagens do conjunto de testes segmentadas com o limiar $p = 0, 10$. Novamente foi observado que o modelo TTA apresentou resultados superiores ao modelo TZ, obtendo 86,3% e 78,2% de $f1-score$ e Intersecção sobre União, respectivamente. Além disso na análise por canal houve um ganho significativo na $f1-score$ em relação ao melhor valor obtido na análise por pixel. Na Figura 23 é mostrado um gráfico com as medidas calculadas para cada valor de p testado.

Na Figura 24 é mostrada uma comparação entre máscaras geradas a partir da marcação manual do especialista e máscaras geradas pelo método com $p=0,1$. Já na Figura 25 são mostradas algumas regiões ampliadas para melhor visualização dos detalhes das marcações.

Na Figura 24 é possível encontrar algumas regiões escuras na matriz óssea indicando que não foram marcados canais nestas regiões. Este é possivelmente um ponto a melhorar

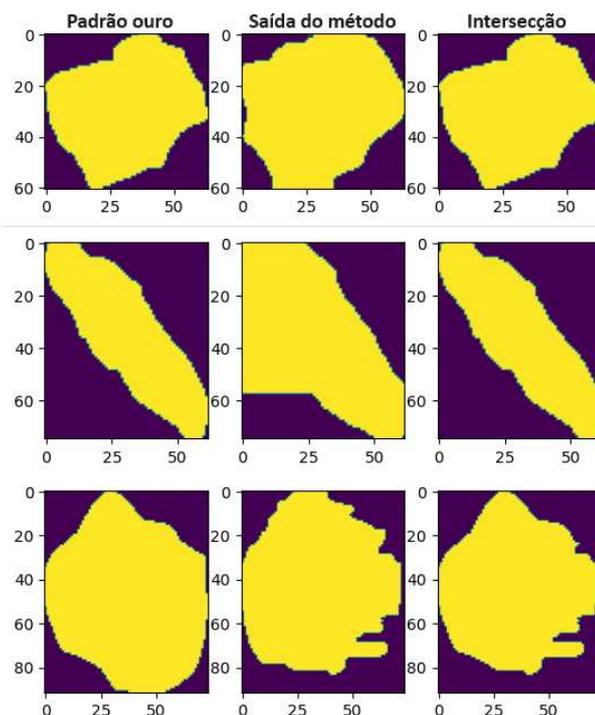


Figura 22 – Três exemplos de componentes conectados analisados durante a análise por canal. Na primeira coluna o componente na máscara de referência em amarelo (padrão ouro). Na segunda coluna a segmentação da rede neural para o respectivo componente. Na terceira coluna a intersecção entre a primeira e a segunda coluna. Fonte: Autor.

no método, e uma das possíveis causas para essas lacunas é o tamanho das sub imagens. Talvez utilizando sub imagens com tamanho 640×640 não tenha sido possível atingir granularidade suficiente para cobrir toda a matriz óssea.

4.3.1 Comparação

A fim de comparação, foi feita uma implementação do algoritmo proposto por (GONDIM et al., 2021) utilizando a linguagem de programação Python na versão 3.10. As imagens de lâmina inteira do conjunto de testes foram então segmentadas e foi realizada a análise por canal conforme descrito anteriormente. Na Figura 26 são mostrados exemplos de componentes conectados obtidos na análise por canal do algoritmo.

Após a contagem de VP, FN e FP realizada na análise por canal, foi observado que no método proposto neste trabalho era menos comum a aparição de falsos negativos. Além disso, algumas imagens do conjunto de testes apresentavam concavidades no formato do corte histológico e o método de (GONDIM et al., 2021) não reagiu bem para tais imagens, conforme mostrado na Figura 27.

Os resultados obtidos em (GONDIM et al., 2021) mostram acurácia e especificidade próximas a 96%, sensibilidade acima de 80% e cerca de 90% de *f1-score*. Porém as

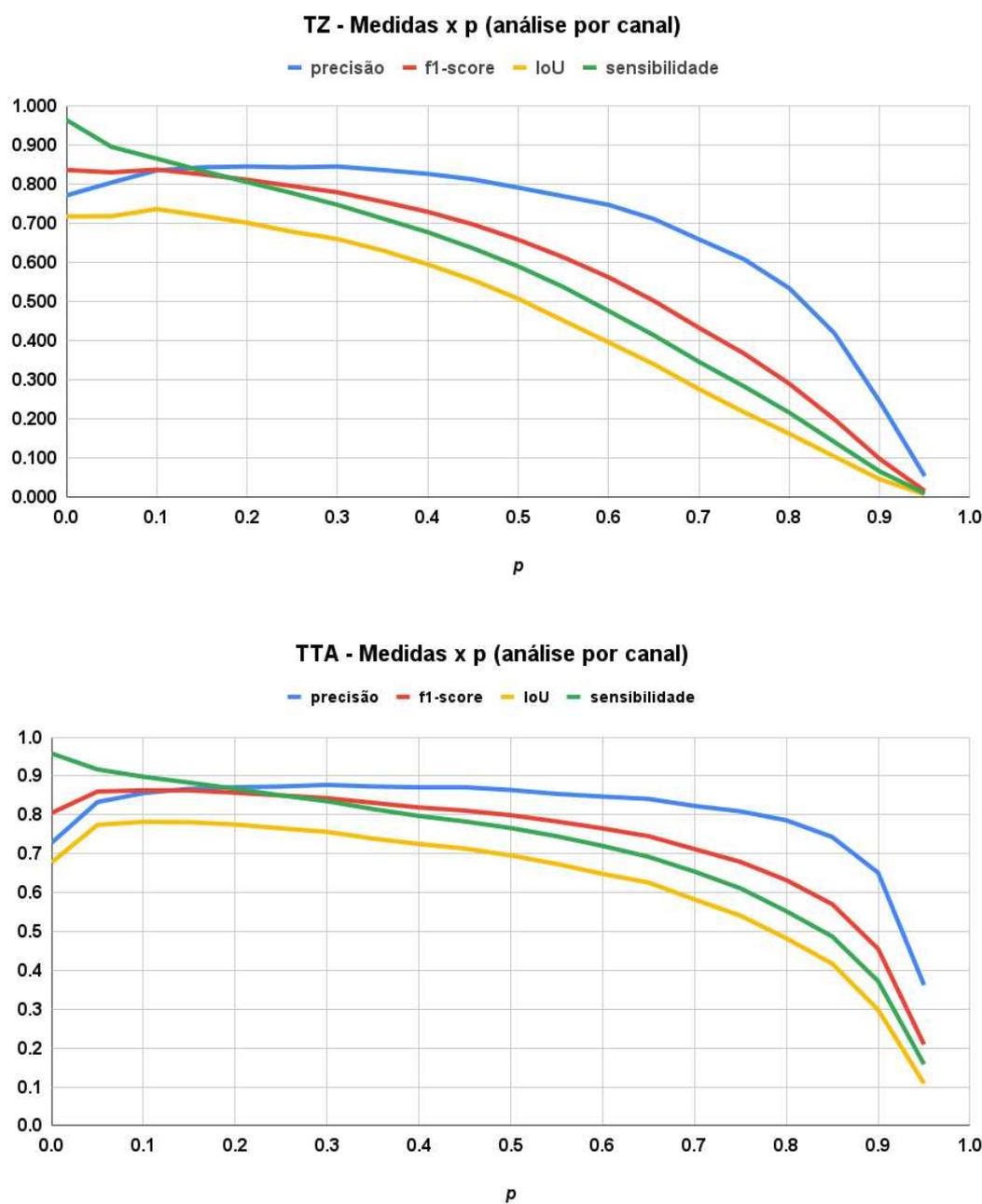


Figura 23 – Precisão, $f1$ -score, Intersecção sobre União e sensibilidade em função do $threshold$ p na análise por canal de ambos os treinamentos. Fonte: Autor.

p	Treinamento do Zero				Transferência de aprendizado			
	Precisão	F1-Score	Sensibilidade	IoU	Precisão	F1-Score	Sensibilidade	IoU
0,05	0,805	0,831	0,896	0,718	0,833	0,860	0,917	0,774
0,10	0,836	0,838	0,866	0,737	0,856	0,863	0,898	0,782
0,15	0,844	0,826	0,835	0,720	0,867	0,863	0,883	0,781
0,20	0,846	0,812	0,806	0,702	0,871	0,857	0,867	0,775
0,25	0,844	0,796	0,778	0,679	0,873	0,850	0,850	0,765
0,30	0,846	0,780	0,748	0,660	0,877	0,843	0,835	0,756
0,35	0,837	0,756	0,713	0,631	0,873	0,831	0,815	0,738
0,40	0,827	0,730	0,678	0,595	0,871	0,819	0,797	0,725
0,45	0,813	0,698	0,637	0,555	0,871	0,811	0,783	0,713
0,50	0,792	0,659	0,591	0,508	0,864	0,799	0,766	0,695
0,55	0,770	0,614	0,538	0,452	0,854	0,783	0,745	0,673
0,60	0,748	0,563	0,477	0,396	0,847	0,765	0,720	0,648
0,65	0,712	0,503	0,415	0,340	0,841	0,745	0,692	0,625
0,70	0,660	0,434	0,347	0,277	0,823	0,712	0,654	0,582
0,75	0,609	0,368	0,284	0,218	0,809	0,679	0,611	0,540
0,80	0,535	0,291	0,217	0,163	0,786	0,632	0,552	0,482
0,85	0,420	0,200	0,142	0,104	0,743	0,570	0,487	0,417
0,90	0,246	0,099	0,067	0,046	0,651	0,455	0,372	0,298
0,95	0,055	0,017	0,011	0,008	0,362	0,209	0,158	0,108

Tabela 2 – Médias de precisão, $f1$ -score, sensibilidade e IoU para cada limiar p testado na análise por canal.

marcações dos especialistas não seguiram a forma exata das estruturas, motivo pelo qual em (GONDIM et al., 2021) foram considerados como verdadeiros positivos os canais cuja marcação do especialista estivesse ao menos 20% contida na marcação feita pelo método, enquanto neste trabalho o critério de aceite foi de 70%.

Não fica claro se as imagens utilizadas em (GONDIM et al., 2021) continham imagens de diferentes regiões da diáfise femoral – região intermediária do fêmur. Devido à anatomia do osso, imagens provenientes da região central da diáfise resultam em cortes histológicos mais arredondados, semelhantes à Figura 27(a), enquanto imagens extraídas de regiões mais próximas às extremidades da diáfise podem resultar em cortes histológicos que apresentem outros formatos, como o observado na Figura 27(c).

A análise por canal – descrita na seção 4.2 – da reprodução do algoritmo proposto por (GONDIM et al., 2021) apresentou um $f1$ -score de aproximadamente 17,5% e uma Intersecção sobre União de 9,6%. A Tabela 3 mostra os resultados obtidos pela análise por canal para cada um dos métodos. Vale ressaltar que esses valores são inferiores aos valores obtidos na metodologia de (GONDIM et al., 2021), cujo critério de aceite era apenas 20%. Já neste trabalho o critério de aceite foi 70% de coincidência dos componentes.

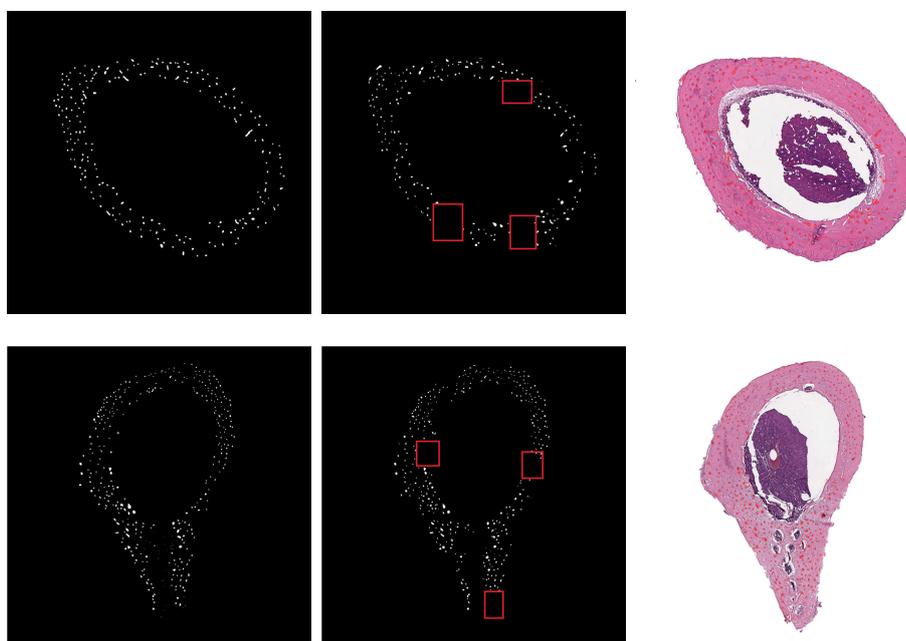


Figura 24 – Comparação entre marcação manual feita por especialista e saída do método. Na primeira coluna exemplos de marcações manuais; na segunda coluna as máscaras geradas pelo método utilizando um limiar $p=0,1$ com algumas lacunas destacadas em vermelho; na terceira coluna a imagem marcada a partir da máscara gerada pela rede. Fonte: Autor.

Método	Precisão	F1-Score	Sensibilidade	IoU
FCN - TZ ($p = 0.10$)	0.826	0.849	0.872	0.737
FCN - TTA ($p = 0.10$)	0.856	0.863	0.898	0.782
Algoritmo (GONDIM et al., 2021)	0.619	0.175	0.102	0.095

Tabela 3 – Médias de precisão, $f1$ -score, sensibilidade e Intersecção sobre União para cada um dos métodos de segmentação testados.

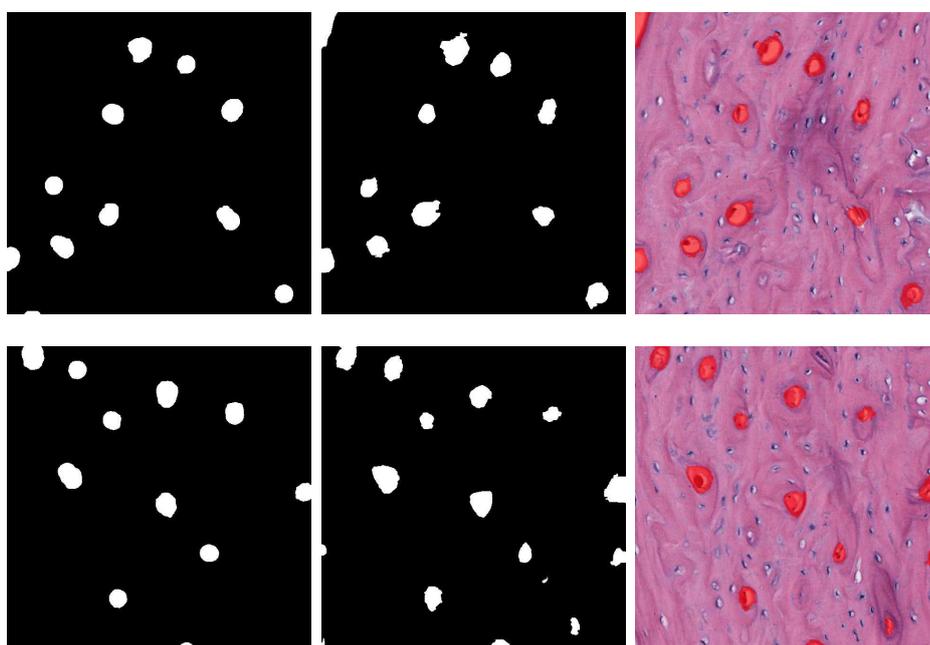


Figura 25 – Comparação de regiões ampliadas das marcações manuais feita por especialista e de saídas do método. Na primeira coluna regiões ampliadas de marcações manuais, na segunda coluna as respectivas regiões nas máscaras geradas pelo método utilizando um limiar $p=0,1$, e na terceira coluna as respectivas regiões na imagens marcadas a partir das máscaras geradas pela rede. Fonte: Autor.

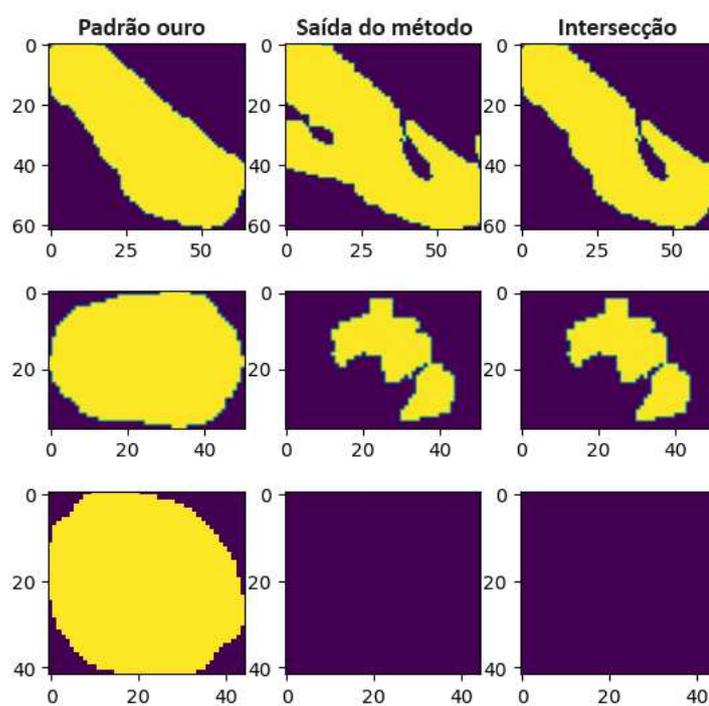


Figura 26 – Três exemplos de componentes conectados analisados. Na primeira coluna o componente na máscara de referência (padrão ouro). Na segunda coluna a segmentação do algoritmo de (GONDIM et al., 2021) para o respectivo componente. Na terceira coluna a intersecção entre a primeira e a segunda coluna. Fonte: Autor.

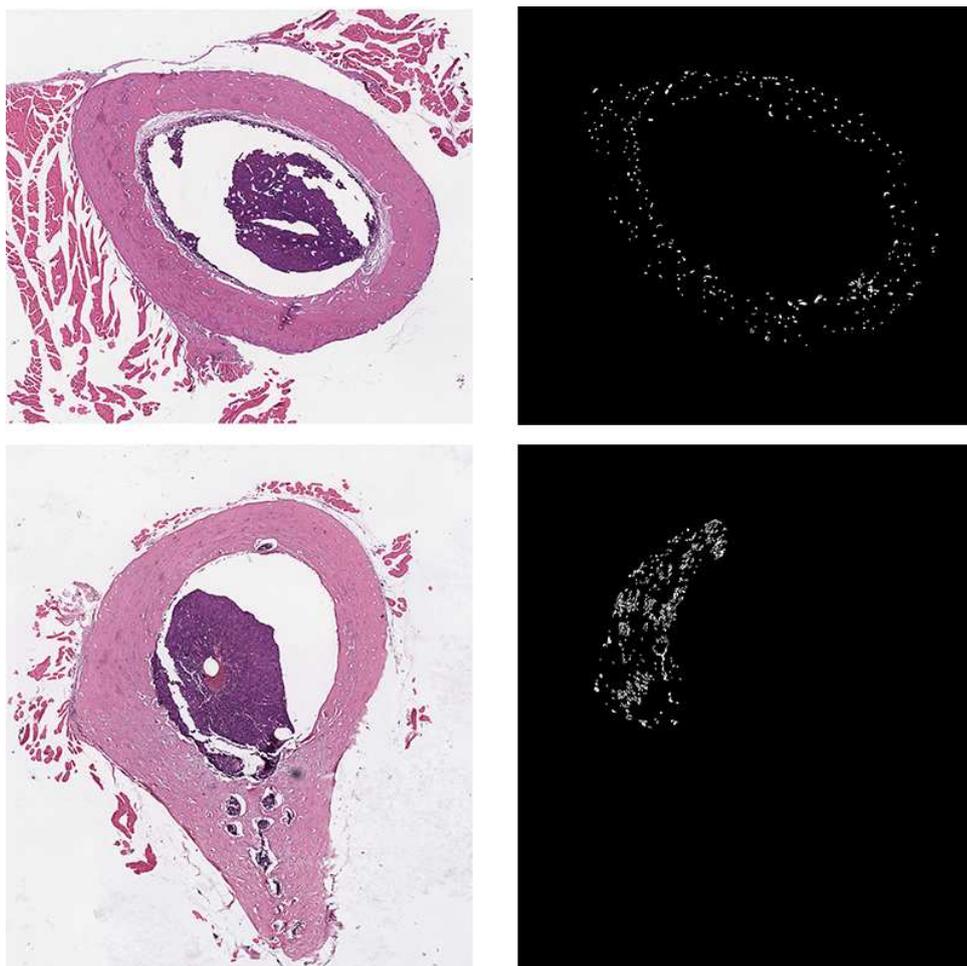


Figura 27 – Exemplo de falha do algoritmo proposto por (GONDIM et al., 2021). Apesar de funcionar para cortes histológicos de formato convexo o método falha ao trabalhar com cortes que apresentem concavidades. Fonte: Autor.

Conclusão

Neste trabalho foi explorado um novo método de segmentação de canais ósseos em imagens histológicas de tecido ósseo utilizando uma rede neural completamente convolucional. A rede foi treinada com um conjunto de imagens próprio, criado a partir de imagens marcadas manualmente com a ajuda de um especialista em histologia. Além disso, o conjunto de imagens utilizado foi disponibilizado publicamente.

O método testado mostrou-se capaz de executar a tarefa de segmentação de canais ósseos em imagens histológicas WSI de tecido ósseo. O uso da técnica de transferência de aprendizado trouxe uma melhora nos resultados obtidos, levando a bons valores de acurácia e especificidade e valores superiores de precisão, sensibilidade e *f1-score*. O método também se mostrou mais robusto e preciso em comparação com outro trabalho encontrado na literatura. Cabe também destacar a qualidade do trabalho apresentado por (SANTOS et al., 2022). Com uma simples adaptação do método por eles proposto, foi possível segmentar um conjunto de imagens de domínio diferente. Isso é um estímulo para o uso da mesma técnica em outros domínios da área da saúde.

Entretanto o trabalho apresenta algumas limitações. Não foram testados outros tamanhos para a divisão da imagem WSI em imagens menores. Uma quebra em tamanhos menores possivelmente melhoraria os resultados visto que poderiam haver mais sub imagens que não apresentam canais. Dessa forma poderíamos descartá-las e treinar a rede com um conjunto de dados mais equilibrado entre *pixels* que são canais e *pixels* que não são canais.

Outra limitação do conjunto de dados foi o fato de não ter sido feita uma análise de concordância sobre as marcações realizadas pelo especialista. Uma análise de concordância com o apoio de um segundo especialista na área seria interessante pois, ao realizar as marcações, o especialista traz consigo um viés de subjetividade. Com a análise de concordância poderíamos ter um conjunto de dados com marcações mais confiáveis, o que traria um resultado mais próximo da realidade. Também não foi validado se o uso de um número maior de épocas durante o treinamento poderia melhorar os resultados.

Conclui-se portanto que a segmentação de canais ósseos em imagens histológicas de

tecido ósseo em imagens de lâmina inteira pode ser realizada utilizando redes neurais completamente convolucionais desde que haja um conjunto de imagens apropriado para o treinamento. Para trabalhos futuros almeja-se aprimorar o método focando nas limitações apresentadas acima. Além disso, pretende-se utilizar o método para segmentar canais ósseos em imagens sequenciais de tecido ósseo, permitindo reconstruir em 3D toda a rede de canais ósseos, tornando viável a realização de análises mais profundas sobre a rede de canais ósseos óssea para pesquisadores da área de histologia.

5.1 Contribuições em Produções Bibliográficas

O artigo *Automatic Segmentation of Bone Canals on Histological Images Using Fully Convolutional Neural Networks* foi submetido para a revista *Biomedical Signal Processing and Control*. Nesse artigo foram reunidas as principais propostas e resultados descritos nesta dissertação.

Outra contribuição deste trabalho, conforme descrito na sub seção 3.1, é a disponibilização pública do conjunto de imagens desenvolvido para que seja utilizado em outros trabalhos.

Além disso, este trabalho foi apresentado na 40^a edição do evento SBPqO (Sociedade Brasileira de Pesquisa Odontológica) em setembro de 2023, levando resultados parciais e submetido para edição 2024 do evento com os resultados finais.

Referências

- BLOCK, H.-D. The perceptron: A model for brain functioning. i. **Reviews of Modern Physics**, APS, v. 34, n. 1, p. 123, 1962. Disponível em: <<https://doi.org/10.1103/RevModPhys.34.123>>.
- CARDOSO, J. S.; CORTE-REAL, L. Toward a generic evaluation of image segmentation. **IEEE Transactions on Image Processing**, IEEE, v. 14, n. 11, p. 1773–1782, 2005. Disponível em: <<https://doi.org/10.1109/TIP.2005.854491>>.
- CHOWDHURY, M. E. H.; RAHMAN, T.; KHANDAKAR, A.; MAZHAR, R.; KADIR, M. A.; MAHBUB, Z. B.; ISLAM, K. R.; KHAN, M. S.; IQBAL, A.; EMADI, N. A.; REAZ, M. B. I.; ISLAM, M. T. Can ai help in screening viral and covid-19 pneumonia? **IEEE Access**, v. 8, p. 132665–132676, 2020. Disponível em: <<https://doi.org/10.1109/ACCESS.2020.3010287>>.
- ESTEVA, A.; CHOU, K.; YEUNG, S.; NAIK, N.; MADANI, A.; MOTTAGHI, A.; LIU, Y.; TOPOL, E.; DEAN, J.; SOCHER, R. Deep learning-enabled medical computer vision. **NPJ digital medicine**, Nature Publishing Group, v. 4, n. 1, p. 1–9, 2021. Disponível em: <<https://doi.org/10.1038/s41746-020-00376-2>>.
- FELDMAN, A. T.; WOLFE, D. Tissue processing and hematoxylin and eosin staining. **Histopathology: megonzalthods and protocols**, Springer, p. 31–43, 2014. Disponível em: <https://doi.org/10.1007/978-1-4939-1050-2_3>.
- GÉRON, A. **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn & TensorFlow**. [S.l.]: Alta Books, 2019.
- GONDIM, P. H. C. C.; LIMIRIO, P. H. J. O.; ROCHA, F. S.; BATISTA, J. D.; DECHICHI, P.; TRAVENÇOLO, B. A. N.; BACKES, A. R. Automatic segmentation of bone canals in histological images. **Journal of Digital Imaging**, Springer, v. 34, n. 3, p. 678–690, 2021. Disponível em: <<https://doi.org/10.1007/s10278-021-00454-1>>.
- GONZÁLEZ-HIDALGO, M.; GUERRERO-PEÑA, F. A.; HEROLD-GARCÍA, S.; CAPÓ, A. Jaume-i; MARRERO-FERNÁNDEZ, P. D. Red blood cell cluster separation from digital images for use in sickle cell disease. **IEEE Journal of Biomedical and Health Informatics**, v. 19, n. 4, p. 1514–1525, 2015. Disponível em: <<https://doi.org/10.1109/JBHI.2014.2356402>>.

GUPTA, V.; DEMIRER, M.; BIGELOW, M.; LITTLE, K. J.; CANDEMIR, S.; PREVEDELLO, L. M.; WHITE, R. D.; O'DONNELL, T. P.; WELS, M.; ERDAL, B. S. Performance of a deep neural network algorithm based on a small medical image dataset: incremental impact of 3d-to-2d reformation combined with novel data augmentation, photometric conversion, or transfer learning. **Journal of digital imaging**, Springer, v. 33, p. 431–438, 2020. Disponível em: <<https://doi.org/10.1007/s10278-019-00267-3>>.

JULIA, R. S. et al. Histological image segmentation of the bone vascular network through deep learning. Universidade Federal de Uberlândia, 2021. Disponível em: <<https://repositorio.ufu.br/handle/123456789/32174>>.

JUNQUEIRA, L. C.; CARNEIRO, J. Histologia básica. In: **Histologia básica**. [S.l.: s.n.], 1985. p. 512–512.

KIRYATI, N.; LANDAU, Y. Dataset growth in medical image analysis research. **Journal of imaging**, MDPI, v. 7, n. 8, p. 155, 2021. Disponível em: <<https://doi.org/10.3390/jimaging7080155>>.

Labati, R. D.; Piuri, V.; Scotti, F. All-idb: The acute lymphoblastic leukemia image database for image processing. In: **2011 18th IEEE International Conference on Image Processing**. [s.n.], 2011. p. 2045–2048. Disponível em: <<https://doi.org/10.1109/ICIP.2011.6115881>>.

LINHARES, C. R. B. et al. Melhoria da regeneração óssea em ratos diabéticos tipo i submetidos à oxigenoterapia hiperbárica. Universidade Federal de Uberlândia, 2019. Disponível em: <<http://doi.org/10.14393/ufu.di.2019.2502>>.

LINHARES, C. R. B.; RABELO, G. D.; LIMIRIO, P. H. J. O.; VENÂNCIO, J. F.; SILVA, I. G. R.; DECHICHI, P. Automated bone healing evaluation: New approach to histomorphometric analysis. **Microscopy Research and Technique**, Wiley Online Library, v. 85, n. 10, p. 3339–3346, 2022. Disponível em: <<https://doi.org/10.1002/jemt.24188>>.

LIU, Z.-Q.; LIEW, H. L.; CLEMENT, J. G.; THOMAS, C. D. L. Bone image segmentation. **IEEE Transactions on biomedical engineering**, IEEE, v. 46, n. 5, p. 565–573, 1999. Disponível em: <<https://doi.org/10.1109/10.759057>>.

LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2015. p. 3431–3440.

MIKLOSIK, A.; EVANS, N. Impact of big data and machine learning on digital transformation in marketing: A literature review. **Ieee Access**, IEEE, v. 8, p. 101284–101292, 2020. Disponível em: <<https://doi.org/10.1109/ACCESS.2020.2998754>>.

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. **Sistemas inteligentes-Fundamentos e aplicações**, Manole, v. 1, n. 1, p. 32, 2003.

MUELLER, J. P.; MASSARON, L. **Deep Learning for dummies**. [S.l.]: John Wiley & Sons, 2019.

PAULLADA, A.; RAJI, I. D.; BENDER, E. M.; DENTON, E.; HANNA, A. Data and its (dis) contents: A survey of dataset development and use in machine learning research. **Patterns**, Elsevier, v. 2, n. 11, p. 100336, 2021.

- RAHMAN, M. A.; WANG, Y. Optimizing intersection-over-union in deep neural networks for image segmentation. In: SPRINGER. **International symposium on visual computing**. 2016. p. 234–244. Disponível em: <https://doi.org/10.1007/978-3-319-50835-1_22>.
- RAWAT, W.; WANG, Z. Deep convolutional neural networks for image classification: A comprehensive review. **Neural computation**, MIT Press, v. 29, n. 9, p. 2352–2449, 2017. Disponível em: <https://doi.org/10.1162/neco_a_00990>.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. **Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18**. 2015. p. 234–241. Disponível em: <https://doi.org/10.1007/978-3-319-24574-4_28>.
- SANTOS, D. F. D. et al. Automated segmentation of tumor regions from oral histological whole slide images using fully convolutional neural networks. Universidade Federal de Uberlândia, 2022. Disponível em: <<http://doi.org/10.14393/ufu.te.2022.252>>.
- SANTOS, D. F. D. dos; FARIA, P. R. de; TRAVENÇOLO, B. A. N.; NASCIMENTO, M. Z. do. Influence of data augmentation strategies on the segmentation of oral histological images using fully convolutional neural networks. **Journal of Digital Imaging**, Springer Science and Business Media LLC, apr 2023. Disponível em: <<https://doi.org/10.1007/s10278-023-00814-z>>.
- SANTOS, D. F. dos; FARIA, P. R. de; TRAVENÇOLO, B. A.; NASCIMENTO, M. Z. do. Automated detection of tumor regions from oral histological whole slide images using fully convolutional neural networks. **Biomedical Signal Processing and Control**, Elsevier BV, v. 69, p. 102921, aug 2021. Disponível em: <<https://doi.org/10.1016/j.bspc.2021.102921>>.
- SIDDIQUE, N.; PAHEDING, S.; ELKIN, C. P.; DEVABHAKTUNI, V. U-net and its variants for medical image segmentation: A review of theory and applications. **Ieee Access**, IEEE, v. 9, p. 82031–82057, 2021. Disponível em: <<https://doi.org/10.1109/ACCESS.2021.3086020>>.
- SIMON, P. **Too Big to Ignore: The Business Case for Big Data**. Wiley, 2013. (Wiley and SAS Business Series). ISBN 9781118642108. Disponível em: <<https://books.google.com.br/books?id=Dn-Gdoh66sgC>>.
- SRINIVASAN, S.; VINAYAKUMAR, R.; ARUNACHALAM, A.; ALAZAB, M.; SOMAN, K. Durl: Malicious url detection using deep learning-based character level representations. In: **Malware analysis using artificial intelligence and deep learning**. Springer, 2021. p. 535–554. Disponível em: <https://doi.org/10.1007/978-3-030-62582-5_21>.
- TORREY, L.; SHAVLIK, J. Transfer learning. In: **Handbook of research on machine learning applications and trends: algorithms, methods, and techniques**. [S.l.]: IGI global, 2010. p. 242–264.
- UDDIN, M. F. Addressing accuracy paradox using enhanced weighted performance metric in machine learning. In: **2019 Sixth HCT Information Technology Trends (ITT)**. [s.n.], 2019. p. 319–324. Disponível em: <<https://doi.org/10.1109/ITT48889.2019.9075071>>.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: SN. **Proceedings of the XXIX conference on graphics, patterns and images**. [S.l.], 2016. v. 1, n. 4.