
**Investigação de técnicas de normalização e
segmentação com características interpretáveis
para imagens histológicas H&E**

André Dias Freitas

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

Uberlândia
2025

André Dias Freitas

**Investigação de técnicas de normalização e
segmentação com características interpretáveis
para imagens histológicas H&E**

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 de Dados

Orientador: Prof. Dr. Marcelo Zanchetta do Nascimento

Coorientador: Prof^ª. Dr^ª. Thaína Aparecida Azevedo Tosta

Uberlândia

2025

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

F866 2025	Freitas, André Dias, 1998- Investigação de técnicas de normalização e segmentação com características interpretáveis para imagens histológicas H&E [recurso eletrônico] / André Dias Freitas. - 2025. Orientador: Marcelo Zanchetta do Nascimento. Coorientadora: Thaína Aparecida Azevedo Tosta. Dissertação (Mestrado) - Universidade Federal de Uberlândia, Pós-graduação em Ciência da Computação. Modo de acesso: Internet. DOI http://doi.org/10.14393/ufu.di.2025.432 Inclui bibliografia. Inclui ilustrações. 1. Computação. I. Nascimento, Marcelo Zanchetta do, 1976-, (Orient.). II. Tosta, Thaína Aparecida Azevedo, 1992-, (Coorient.). III. Universidade Federal de Uberlândia. Pós-graduação em Ciência da Computação. IV. Título. CDU: 681.3
--------------	--

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:
Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em 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, 08/2025, PPGCO				
Data:	11 de Julho de 2025	Hora de início:	09:05	Hora de encerramento:	10:50
Matrícula do Discente:	12222CCP003				
Nome do Discente:	André Dias Freitas				
Título do Trabalho:	Investigação de técnicas de normalização e segmentação com características interpretáveis para imagens histológicas H&E				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Ciência de Dados				
Projeto de Pesquisa de vinculação:	APQ-00727-24 - Ensemble Learning e Geometria Fractal Empregados para Classificação de Tumores Odontogênicos Bucais em Imagens Histológicas HE - FAPEMIG. 302833/2025-0 - Desenvolvimento de Modelos Híbridos para Segmentação e Classificação de Imagens da Cavidade Oral - CNPq. 2022/03020-1 - Normalização de corantes H&E por autocodificadores com análises de ensemble learning para imagens histológicas - FAPESP				

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Thaína Aparecida Azevedo Tosta - UNIFESP (Coorientadora), Henrique Coelho Fernandes - Cranfield University, Marcia Aparecida Silva Bissaco - UMC e Marcelo Zanchetta do Nascimento - FACOM/UFU, orientador do candidato.

Os examinadores participaram desde as seguintes localidades: Thaína Aparecida Azevedo Tosta - São José dos Campos/SP, Henrique Coelho Ferenandes - Milton Keynes/Inglaterra e Marcia Aparecida Silva Bissaco -Mogi das Cruzes/SP. O Marcelo Zanchetta do Nascimento e o aluno participaram da cidade de Uberlândia.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Marcelo Zanchetta do Nascimento , 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 da 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 **Marcia Aparecida Silva Bissaco, Usuário Externo**, em 14/07/2025, às 14:50, 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 **Henrique Coelho Fernandes, Professor(a) do Magistério Superior**, em 14/07/2025, às 16:00, 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 **Thaína Aparecida Azevedo Tosta, Usuário Externo**, em 14/07/2025, às 17:31, 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 14/07/2025, às 19:37, 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 **6488739** e o código CRC **4F3EF598**.

Dedico este trabalho a minha família por todo suporte, e aos meus orientadores que me auxiliaram durante todo o processo.

Agradecimentos

Agradeço aos meus orientadores, Marcelo Zanchetta do Nascimento e Thaína Aparecida Azevedo Tosta, pela paciência na orientação, pelo convívio, pelo apoio, pela compreensão e pela amizade que tornaram possível a conclusão deste projeto. Agradeço também a André Fernando Quaresma da Silva e Adriano Barbosa Silva, que me auxiliaram em pontos específicos deste trabalho.

O orientador deste projeto agradece o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (Projeto: 302833/2025-0) e da Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG (Projeto: APQ-00727-24) e Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP (Projeto: 2022/03020-1).

“A vida é como andar de bicicleta. Para manter o equilíbrio, você precisa continuar em movimento.”
(Albert Einstein)

Resumo

A displasia epitelial oral é uma condição pré-cancerosa que pode ser identificada por meio da análise de imagens histológicas, geralmente realizada por especialistas. No entanto, o alto volume de trabalho e a variabilidade na experiência clínica tornam esse processo sujeito à subjetividade. Como forma de apoio, ferramentas computacionais têm sido utilizadas com leitura suplementar para auxiliar na tomada de decisão. Este trabalho apresenta uma investigação entre diferentes métodos de normalização de corantes de Hematoxilina e Eosina (H&E) em imagens histológicas e o desenvolvimento de um segmentador de núcleos baseado em uma arquitetura da U-net com interpretabilidade das características. Foram testados métodos de normalização de corantes baseados em aprendizado de dicionário, os quais foram avaliados com quatro métricas distintas. Esses dados foram avaliados em uma arquitetura de segmentação baseada em uma U-Net modificada. Foi explorada a interpretabilidade dos descritores nas camadas internas do modelo com o objetivo de avaliar a influência da normalização e o melhoramento do desempenho da segmentação com a aplicação da técnica ProtoSeg. Os resultados mostraram que o método de normalização baseado em Distância Euclidiana proporcionou valores médios de 0,897, 44,401, 0,898 e 0,973, respectivamente, para as métricas FSIM, PSNR, QSSIM e SSIM. Além disso, os resultados mostraram que há métodos de normalização que melhoraram a etapa de segmentação, alcançando valores de coeficiente *Dice* em torno de 0,7683. No entanto, quando aplicada a estratégia ProtoSeg em conjunto com a normalização, esses resultados não foram superiores. Esse estudo mostrou que a normalização de corantes H&E teve impacto positivo na segmentação apenas quando não é aplicada a interpretabilidade na segmentação.

Palavras-chave: Displasia epitelial oral. Imagens Histológicas. Segmentação de Imagens. Normalização de Cores. Redes Neurais Convolucionais. Interpretabilidade.

Abstract

Oral epithelial dysplasia is a precancerous condition that can be identified through the analysis of histological images, typically performed by specialists. However, the high workload and variability in clinical experience make this process susceptible to subjectivity. As a form of support, computational tools have been used as a supplementary aid in decision-making. This work presents an investigation into different Hematoxylin and Eosin (H&E) stain normalization methods in histological images and the development of a nuclei segmenter based on a U-Net architecture with feature interpretability. Dictionary learning-based stain normalization methods were tested and evaluated using four different metrics. These normalized datasets were assessed using a modified U-Net segmentation architecture. Interpretability of the descriptors in the internal layers of the model was explored to evaluate the influence of normalization and the enhancement of segmentation performance through the application of the ProtoSeg technique. The results showed that the normalization method based on Euclidean Distance yielded average values of 0.897, 44.401, 0.898, and 0.973 for the FSIM, PSNR, QSSIM, and SSIM metrics, respectively. Furthermore, the findings indicated that certain normalization methods improved the segmentation stage, reaching *Dice* coefficient values around 0.7683. However, when the ProtoSeg strategy was applied alongside normalization, these results were not superior. This study demonstrated that H&E stain normalization had a positive impact on segmentation performance only when interpretability was not applied during the segmentation process.

Keywords: Oral Epithelial Dysplasia. Histological Images. Image Segmentation. Color Normalization. Convolutional Neural Networks. Interpretability.

Lista de ilustrações

Figura 1	– Exemplos de tecidos histológicos orais nos diferentes níveis de displasia: (a) Saudável, (b) Leve, (c) Moderada e (d) Severa. As setas vermelhas indicam os núcleos celulares e as azuis indicam o epitélio (Figura extraída de (A. B. Silva 2019)).	34
Figura 2	– Principais etapas para análise de imagens histológicas (Figura adaptada de (D. F. D. d. Santos 2022)).	35
Figura 3	– Etapas e exemplos de técnicas realizadas em um sistema <i>Computer-Aided Diagnosis</i> (CAD) (Figura adaptada de (W. He et al. 2022)). . .	37
Figura 4	– Exemplo de normalização de uma imagem original (b), dada uma imagem de referência (a), originando uma imagem normalizada (c) (Figura extraída de (TOSTA et al. 2019a)).	38
Figura 5	– Representação gráfica dos vetores de coloração efetivos (Figura adaptada de (Vahadane et al. 2016)).	42
Figura 6	– Uma arquitetura simples de rede neural artificial (Figura adaptada de (Rosebrock 2017)).	43
Figura 7	– Funcionamento de um <i>kernel</i> sobre uma região da imagem (Figura extraída de (Rosebrock 2017)).	43
Figura 8	– Exemplos de transformações por diferentes tipos de <i>kernels</i> (Figura extraída de (Rosebrock 2017)).	44
Figura 9	– Exemplos das operações <i>max pooling</i> (b) e <i>average pooling</i> (c) em uma matriz de tamanho 4×4 pixels (a), utilizando <i>pool</i> de tamanho 2×2 pixels (Figura extraída de (D. F. D. d. Santos 2022)).	45
Figura 10	– Arquitetura do modelo <i>LeNet</i> . (Figura extraída de (Lecun et al. 1998)).	45
Figura 11	– Arquitetura do modelo <i>AlexNet</i> . (Figura extraída de (Krizhevsky et al. 2017)).	46
Figura 12	– Exemplo das principais tarefas de visão computacional (Figura adaptada de (Hao et al. 2020)).	46

Figura 13 – Exemplo de arquitetura de uma rede U-Net (Figura extraída de (Ronneberger et al. 2015)).	48
Figura 14 – Diagrama das etapas da metodologia proposta para normalização, segmentação e interpretação das características.	58
Figura 15 – Tecidos histológicos da cavidade oral com imagem saudável e com lesões displásicas. Na linha superior são mostradas as <i>Region of Interest</i> (ROI)s e a linha inferior apresenta as máscaras do padrão ouro delineado pelo especialista (Figura adaptada de (A. B. Silva et al. 2024b)).	58
Figura 16 – Exemplo de <i>crop</i> realizado na base de imagens para obter imagens de entrada no padrão da arquitetura usada para segmentação.	59
Figura 17 – Gráfico de treinamento e validação de uma rede CNN com <i>overfitting</i> . .	62
Figura 18 – Arquitetura da rede CNN apresentada neste trabalho.	63
Figura 19 – Gráficos de treinamento utilizando os diferentes abordagens de regularização analisadas no modelo.	64
Figura 20 – Estrutura geral da metodologia ProtoSeg para geração do <i>Segmentation Ability Score</i> (<i>SA Score</i>).	67
Figura 21 – Imagens de referência utilizadas no processo de normalização dos corantes Hematoxilina e Eosina (H&E).	72
Figura 22 – Resultados da normalização de H&E usando a imagem Ref 1 para investigação dos métodos.	73
Figura 23 – Exemplos de normalizações de cores realizadas com cada método utilizando a imagem Ref 2.	73
Figura 24 – Investigação da normalização dos corantes realizada com cada método utilizando a imagem Ref 3.	74
Figura 25 – Exemplos de resultados da segmentação de núcleos celulares, comparando a máscara padrão ouro e os resultados do método KL utilizando a imagem Ref 2. Os FP são representados pela cor verde e os FN pela cor vermelha.	78
Figura 26 – Gráficos de treinamento do modelo com imagens H&E sem normalização: (esquerda) curva da acurácia versus épocas; (direita) taxa de perda na etapa de treinamento e validação.	79
Figura 27 – Gráficos de treinamento da rede com imagens normalizadas com o método KL a partir da imagem Ref 2: (esquerda) curva da acurácia versus épocas; (direita) taxa de perda na etapa de treinamento e validação. .	79
Figura 28 – Valores médios e desvios padrões dos <i>SA Scores</i> de cada camada do modelo com imagens não normalizadas com e sem uso da estratégia online do Protoseg.	80

Figura 29 – Resultados com as médias e desvios padrões dos <i>SA Scores</i> calculados para cada camadas da rede treinada com imagens normalizadas pelo método KL a partir da imagem Ref 2.	80
Figura 30 – Gráficos de treinamento da rede treinada com imagens sem normalização com o método Protoseg: (esquerda) curva da acurácia versus épocas; (direita) taxa de perda na etapa de treinamento e validação. . .	82
Figura 31 – Gráficos de treinamento da rede treinada com imagens normalizadas com o método KL a partir da imagem Ref 2 e com o método Protoseg Online: (esquerda) curva da acurácia versus épocas; (direita) taxa de perda na etapa de treinamento e validação.	82

Lista de tabelas

Tabela 1	– Métricas de normalização dos métodos estudados com uso da imagem Ref 1.	74
Tabela 2	– Resultados quantitativos das técnicas de normalização pela imagem Ref 2.	75
Tabela 3	– Métricas de normalização dos métodos estudados, a partir da imagem de Ref 3.	75
Tabela 4	– Valores das métricas Acurácia e coeficiente <i>Dice</i> obtidas na segmenta- ção extraídas na etapa de teste do modelo.	76
Tabela 5	– Resultados com as métricas de segmentação com o método de segmen- tação e a abordagem Protoseg.	81
Tabela 6	– Avaliação da generalização do modelo e influência da normalização so- bre os dados.	83
Tabela 7	– Avaliação entre as diferentes abordagens para segmentação de núcleos em imagens histológicas de displasia.	83

Lista de siglas

ANN *Artificial Neural Networks*

BKSVD *Bayesian K-SVD*

CAD *Computer-Aided Diagnosis*

CNN *Convolutional Neural Networks*

DEO *Displasia Epitelial Oral*

EBKSVD *Empirical Bayesian K-SVD*

ED *Euclidean Distance*

FN *Falso Negativo*

FP *Falso Positivo*

FC *Fully-Connected*

FSIM *Feature Similarity Indexing Method*

H&E Hematoxilina e Eosina

hpcNMF High-Performance Computing Non-negative Matrix Factorization

K-SVD *K-Singular Value Decomposition*

KL *Kullback-Leibler*

NMF *Non-negative Matrix Factorization*

OMS Organização Mundial da Saúde

PSNR *Peak Signal-to-Noise Ratio*

QSSIM *Quaternion Structural SIMilarity*

ReLU *Rectified Linear Unit*

ROI *Region of Interest*

RGB *Red, Green and Blue*

RRMSE *Relative Root Mean Squared Error*

SNMF *Sparse Non-negative Matrix Factorization*

SVD *Singular Value Decomposition*

SSIM *Structural Similarity Index Measure*

SAM *Segmentation Ability Map*

SA Score *Segmentation Ability Score*

VP Verdadeiro Positivo

VN Verdadeiro Negativo

Sumário

1	INTRODUÇÃO	27
1.1	Motivação	29
1.2	Objetivos e Desafios da Pesquisa	30
1.3	Hipóteses	30
1.4	Contribuições	31
1.5	Organização da Dissertação	31
2	FUNDAMENTAÇÃO TEÓRICA	33
2.1	Displasias Epiteliais Orais	33
2.2	Imagens Histológicas com Corantes H&E	35
2.3	Sistemas de Apoio a Diagnósticos	36
2.4	Normalização de Imagens com Corantes H&E	38
2.5	Métodos de Aprendizado de Dicionários	40
2.6	Redes Neurais Convolucionais	41
2.7	Segmentação de Imagens	46
2.8	Explicabilidade de Redes Neurais Artificiais	48
2.9	Considerações Finais	49
3	TRABALHOS RELACIONADOS: ESTADO DA ARTE	51
3.1	Normalização de Corantes em Imagens H&E	51
3.2	Segmentação de Núcleos em Imagens H&E de Cavidade Oral	54
3.3	Considerações finais	56
4	METODOLOGIA	57
4.1	Introdução	57
4.2	Banco de Imagens	57
4.3	Normalização de Corantes de Hematoxilina e Eosina	59
4.4	Estratégia de Aumento de Dados	62

4.5	Segmentação de Núcleos Celulares nas Imagens Histológicas .	63
4.6	Interpretação das Características no Modelo de Segmentação .	65
4.7	Métricas e critérios de avaliação	67
5	EXPERIMENTOS E ANÁLISE DOS RESULTADOS	71
5.1	Métodos de normalização de corantes baseados em aprendi- zado de dicionário	71
5.2	Segmentação de núcleos celulares em imagens histológicas de displasia	76
5.3	Interpretação das características na segmentação das imagens .	78
5.4	Exploração do Conjunto de Dados com os Modelos de Apre- ndizagem Profunda	82
5.5	Avaliação com o Estado da Arte	83
6	CONCLUSÃO	85
6.1	Principais Contribuições	86
6.2	Trabalhos Futuros	87
6.3	Contribuições em Produção Bibliográfica	87
	Bibliografia	89

Introdução

O câncer é a segunda causa mais comum de óbito conforme dados da Organização Mundial da Saúde, com uma crescente taxa de incidência estimada de 20 milhões de novos casos até 2025 (Organização Mundial da Saúde 2020). Apontado como um dos principais problemas de saúde pública, o câncer ocupa uma posição significativa como uma das principais causas de morte prematura em diversos países, atrás apenas de doenças cardiovasculares (Wild et al. 2020). De acordo com o Instituto Nacional de Câncer (INCA), estima-se que ocorreram anualmente 704 mil novos casos de câncer no Brasil no período de 2023 a 2025, sendo 15.100 casos de câncer da cavidade oral. Esse tipo de câncer ocupa o oitavo lugar em termos de mortalidade global (M. d. O. Santos et al. 2023). Esses dados destacam a magnitude do câncer como um problema de saúde pública e a necessidade de esforços contínuos na prevenção, diagnóstico precoce e tratamento dessa doença. Pesquisas relacionadas a esse tema contribuem para a compreensão e controle do câncer, sendo essenciais para evitar riscos à saúde e tomar medidas preventivas adequadas (Oliveira Santos 2020).

Uma maneira de contribuir para a redução do número de casos da doença é por meio do diagnóstico de lesões pré-cancerosas, aquelas com potencial evolução maligna (Tilakaratne et al. 2019). Nesse contexto, a identificação precoce e a categorização da gravidade dessas lesões são de extrema importância para a saúde do paciente, pois permitem a escolha de um tratamento adequado (V. Kumar et al. 2010; Oliveira Santos 2020). No entanto, o diagnóstico dessas lesões é influenciado por fatores como a carga de trabalho e a subjetividade de especialistas, o que pode levar a divergências nos diagnósticos (Smith et al. 2009). Isso ocorre devido ao fato de que o diagnóstico é baseado no tamanho da lesão e no grau de alterações nucleares, características que não estão necessariamente relacionadas. Portanto, é possível que um especialista atribua o mesmo grau de gravidade a lesões de diferentes severidades (Warnakulasuriya et al. 2008; V. Kumar et al. 2010).

O diagnóstico dessas lesões é comumente realizado por meio da análise microscópica de tecidos, que permite a tomada de decisões relacionadas ao tratamento do paciente (V. Kumar et al. 2010). Nesse processo, uma etapa crucial é a identificação do epitélio,

região onde ocorre o início da manifestação dessas lesões. Além disso, a análise histológica realizada por especialistas envolve a identificação e caracterização dos núcleos que compõem o tecido, pois as alterações morfológicas dessas estruturas são utilizadas para determinar o nível de gravidade da lesão (Müller 2018). No entanto, essa análise é uma tarefa complexa que demanda tempo e pode ser influenciada pela subjetividade do especialista. Estudos demonstraram que a dupla leitura, isto é, uma análise realizada por dois especialistas diferentes, pode ajudar a minimizar tais efeitos durante a análise (Brennan et al. 2019; Zelst et al. 2020). Com os avanços tecnológicos, os sistemas computacionais de apoio a diagnósticos (do inglês, CAD) têm sido adotados como uma solução para auxiliar especialistas, servindo como uma segunda opinião para a definição do diagnóstico.

Os sistemas CAD têm sido propostos como ferramentas auxiliares para patologistas, pois realizam análises quantitativas das informações histológicas (Chan et al. 2020a). Dentre as técnicas empregadas no diagnóstico de anomalias em sistemas CAD, destaca-se o processamento digital de imagens, que permite a quantificação das estruturas do tecido e a análise detalhada de suas características (N. Kumar et al. 2019; Minaee et al. 2021). Essa abordagem quantitativa visa extrair representações numéricas das estruturas histológicas, fornecendo informações relevantes para o diagnóstico do tipo de lesão presente em uma amostra de tecido. Esse sistema abrange várias etapas, incluindo pré-processamento, segmentação, extração de características, classificação, visualização e validação (Yanase et al. 2019). Na etapa de pré-processamento, a imagem é preparada para a aplicação dos algoritmos de análise, envolvendo operações e filtros para redução ou remoção de ruídos, normalização de brilho e cores, bem como o realce de regiões ou objetos relevantes para a análise, por segmentação (Yanase et al. 2019). Dentre esses processos, as etapas de normalização e segmentação são especialmente críticas, pois influenciam diretamente a qualidade dos dados analisados nas fases subsequentes.

Durante o pré-processamento, a imagem é preparada para análise em etapas subsequentes com emprego de operações que visam reduzir ruídos, padronizar características visuais e realçar estruturas relevantes. A normalização de corantes H&E desempenha papel central ao corrigir variações introduzidas por diferentes protocolos de coloração. Técnicas como a normalização de corantes garantem uma uniformidade cromática entre as amostras, contribuindo para a robustez dos algoritmos de segmentação e classificação (Belsare et al. 2016).

A segmentação é uma etapa fundamental na análise de imagens médicas, pois consiste na divisão da imagem em regiões com base em critérios específicos, permitindo a remoção de áreas irrelevantes e concentrando a análise nos objetos de interesse, como os núcleos celulares (Yang et al. 2017). Tradicionalmente, esse processo é realizado por meio de técnicas como limiarização, detecção de bordas e agrupamento de regiões similares. No entanto, abordagens baseadas em aprendizado de máquina têm ganhado destaque por sua capacidade de lidar com maior complexidade e variabilidade nos dados (El-Dahshan et al.

2014; Liu et al. 2018). Com a imagem segmentada, é possível extrair informações numéricas que representam essas regiões, convertendo dados visuais em atributos úteis para os algoritmos de classificação. Essas características variam desde informações morfológicas até propriedades de textura, sendo que a precisão da segmentação afeta diretamente a qualidade dos atributos extraídos. No entanto, segmentar imagens histológicas apresenta desafios relevantes, como a variabilidade na aparência celular, bordas indefinidas e sobreposição ou aglomeração de células (Xie et al. 2019). Ainda, os artefatos e ruídos introduzidos durante a preparação das lâminas histológicas podem interferir no processo, exigindo métodos mais robustos e adaptativos (Basu et al. 2024). A qualidade da segmentação impacta diretamente a extração de características, que são essenciais para a classificação precisa das imagens (Wei et al. 2017; Humeau-Heurtier 2019).

O desempenho da etapa de segmentação depende do contraste visual entre núcleos e citoplasma, um fator que pode ser prejudicado por inconsistências na coloração obtidas pelos corantes. Nesse sentido, a normalização da imagem antes da segmentação não apenas pode melhorar a qualidade visual das amostras, mas também aumentar a confiabilidade dos algoritmos segmentadores, reduzindo erros como a inclusão de regiões de fundo ou a perda de núcleos mal destacados. A integração de técnicas eficazes de normalização e segmentação com foco em características visuais e biológicas é essencial para o sucesso de um sistema CAD aplicado a imagens histológicas H&E. A padronização cromática contribui para maior consistência nos dados, enquanto a segmentação precisa viabiliza a extração de atributos significativos para a classificação automatizada (Hameed et al. 2022). Com isso, investigá-las em conjunto é crucial para o desenvolvimento de modelos mais robustos, interpretáveis e confiáveis no apoio ao diagnóstico patológico em sistemas CAD.

1.1 Motivação

O diagnóstico precoce e preciso, juntamente com a correta determinação do grau de severidade da doença, são importantes ao permitir a administração adequada do tratamento antes que a lesão se agrave ou evolua para câncer, representando riscos à vida do paciente (V. Kumar et al. 2013). O procedimento manual de diagnóstico de lesões envolve etapas complexas e demoradas, o que requer um tempo considerável do patologista e está sujeito à experiência desse profissional (Kademani 2019).

No contexto médico, técnicas computacionais têm sido empregadas devido à sua capacidade de extrair e analisar informações de forma rápida. Essas técnicas possibilitam a realização de análises quantitativas com base em informações e padrões, superando as limitações do processo manual (Solomon et al. 2000). Portanto, o desenvolvimento de novas ferramentas computacionais para auxiliar na identificação de lesões cancerosas é de grande relevância.

O processo de normalização de cores dos corantes H&E exerce um papel importante na aplicação e otimização dessas ferramentas. A variação das características e qualidades dos corantes utilizados é algo importante a ser considerado, e ainda são percebidas discrepâncias entre os estudos da área, em que alguns métodos são definidos como úteis em alguns trabalhos e com desempenho limitado em outros (TOSTA et al. 2019b). Além disso, não existe uma definição precisa de como a normalização de cores afeta o processo de segmentação das imagens, e mais especificamente, como os diferentes métodos de normalização podem afetar esse processo. Esse cenário motiva este trabalho de pesquisa a propor um novo estudo comparativo entre métodos de normalização e uma análise aprofundada em como essas técnicas podem afetar a etapa de segmentação de imagens.

1.2 Objetivos e Desafios da Pesquisa

Este trabalho propõe uma abordagem para a segmentação e interpretação de características em imagens histológicas coradas com H&E, previamente normalizadas. Além disso, é realizado um estudo comparativo entre diferentes técnicas de normalização de corantes baseadas em aprendizado de dicionário, considerando a variação na escolha das imagens de referência. Portanto, pretende-se pontualmente:

- a) Realizar um estudo de abordagens de aprendizado de dicionário para a normalização de cores H&E em imagens de displasia;
- b) Desenvolver um modelo baseado na abordagem U-Net para segmentação dos núcleos celulares nas imagens histológicas para as displasias epiteliais orais;
- c) Investigar os impactos da normalização de cores na metodologia proposta, realizando uma comparação entre os diferentes métodos de normalização utilizados;
- d) Analisar a abordagem ProtoSeg para investigação das características interpretáveis na segmentação com as imagens normalizadas;
- e) Avaliar comparativamente as bases de imagens (originais e normalizadas) para análise de generalização e avaliação do impacto da normalização no modelo desenvolvido;
- f) Analisar o desempenho da metodologia proposta em comparação com trabalhos relacionados na literatura por meio de métricas de desempenho.

1.3 Hipóteses

As técnicas de normalização para imagens histológicas têm-se mostrado promissoras em estudos voltados à análise de lesões histológicas, especialmente no que diz respeito à redução da não uniformidade causada por diferentes protocolos de coloração (Sha et al. 2017; Pérez-Bueno et al. 2022). A Hematoxilina e Eosina H&E, corante padrão em

histopatologia, pode apresentar variações significativas de coloração entre lâminas, o que impacta negativamente o desempenho de sistemas CAD (Pérez-Bueno et al. 2022).

Embora diversos métodos de segmentação de núcleos tenham sido propostos na literatura, poucos incorporam explicitamente etapas de normalização de cor em seus fluxos de processamento. Ainda observa-se um crescente interesse na utilização de métodos de interpretação no contexto de redes neurais profundas, visando tornar os modelos mais compreensíveis. Essas abordagens podem colaborar para uma segmentação mais precisa e confiável. Diante desse contexto, esta pesquisa parte das hipóteses:

- ❑ H1: A utilização de técnicas para normalização dos corantes H&E baseadas em aprendizado de dicionário proporciona melhorias significativas nos resultados de segmentação, em comparação com o uso de imagens não normalizadas;
- ❑ H2: A incorporação de métodos de interpretação no processo de segmentação permite o desenvolvimento de redes neurais com desempenho superior, ao evidenciar características relevantes, mesmo em casos com artefatos.

1.4 Contribuições

As principais contribuições deste trabalho são:

- ❑ Uma avaliação comparativa da eficiência de diferentes métodos de aprendizado de dicionários aplicados à normalização de corantes H&E em imagens histológicas;
- ❑ Desenvolvimento de uma arquitetura otimizada de rede neural profunda baseada em U-Net para a segmentação de núcleos celulares em imagens histológicas H&E;
- ❑ Uma análise do impacto dos diferentes métodos de normalização no desempenho do modelo de segmentação, considerando sua aplicação tanto no treinamento quanto no teste;
- ❑ Um estudo sobre a utilização de técnicas de interpretação ProtoSeg e seus efeitos no processo de treinamento e nos resultados do modelo de segmentação.

1.5 Organização da Dissertação

Este capítulo apresentou os objetivos da metodologia proposta e a motivação para o desenvolvimento deste trabalho. Os próximos capítulos são organizados da seguinte maneira:

- ❑ **Capítulo 2 - Fundamentação Teórica:** apresenta conceitos de lesões pré-câncer e de processamento de imagens importantes para compreender este trabalho, como normalização de imagens, redes neurais artificiais e segmentação de imagens;

- ❑ **Capítulo 3 - Trabalhos Relacionados: Estado da Arte:** apresenta o estado da arte para processamento digital de imagens de diferentes lesões cancerosas. São detalhados trabalhos focados em normalização de cores em imagens coradas com H&E, e trabalhos dedicados à segmentação de núcleos celulares em imagens de lesões cancerosas da região da cavidade oral;
- ❑ **Capítulo 4 - Metodologia:** descreve o banco de imagens de displasia, descreve as técnicas empregadas nas etapas de normalização de imagens, aumento de dados e o modelo de CNN investigado para segmentação das imagens, bem como as métricas utilizadas para a avaliação do método, além de apresentar alguns conceitos específicos utilizados pela metodologia;
- ❑ **Capítulo 5 - Experimentos e Análise dos Resultados:** são apresentados os resultados obtidos em cada etapa da metodologia, resultados empíricos e os obtidos por outros métodos presentes na literatura. São apresentadas discussões e comparações entre o método analisado e os demais;
- ❑ **Capítulo 6 - Conclusão:** resume os principais pontos do trabalho, destaca o que foi alcançado, reconhece as limitações encontradas e sugere caminhos para estudos futuros.

Fundamentação Teórica

Neste capítulo, são apresentados conceitos fundamentais do tema de pesquisa para a compreensão deste trabalho. Serão explorados temas provenientes de diferentes áreas, incluindo ciência da computação e biomedicina.

2.1 Displasias Epiteliais Orais

A Displasia Epitelial Oral (DEO) é considerada uma lesão potencialmente maligna, caracterizada por alterações citológicas e desorganização na maturação celular (Kademani 2019). Essas mudanças afetam tanto a morfologia das células quanto a estrutura do tecido epitelial, sendo classificadas como lesões pré-cancerosas quando ocorrem no epitélio da cavidade oral. Entre as alterações celulares observadas estão a hipercromasia (cor mais intensa ou escura do que o normal no núcleo de uma célula), o crescimento anormal, a perda de uniformidade e deformações nucleares (V. Kumar et al. 2010). Além disso, a organização estrutural do epitélio também pode ser afetada. Em vez de apresentar a progressão ordenada das células basais até as camadas superficiais, observa-se, em alguns casos, a presença de células basais ao longo de toda a espessura do tecido, indicando uma perda do padrão normal de diferenciação celular (V. Kumar et al. 2013).

O diagnóstico e a classificação da DEO seguem as diretrizes estabelecidas pela Organização Mundial da Saúde (OMS), publicadas em 2005, que consideram tanto a gravidade das alterações celulares quanto a extensão do epitélio afetado (Kademani 2019). Lesões leves são aquelas restritas ao terço basal do epitélio, com alterações discretas. Já as formas severas comprometem mais de dois terços da espessura epitelial e apresentam alterações celulares acentuadas (Warnakulasuriya et al. 2008). Entre esses extremos estão as lesões moderadas, que afetam de um a dois terços do epitélio e exigem avaliação cuidadosa. Dependendo da intensidade das alterações, podem ser reclassificadas como leves ou severas (Smith et al. 2009).

Na Figura 1, são representados tecidos obtidos em diferentes níveis de severidade de displasia. A displasia é uma forma de lesão que se caracteriza pela alteração morfológica

das células, incluindo mudanças no tamanho, formato e características internas (A. B. Silva 2019). As imagens presentes na Figura 1 foram coradas com H&E, onde as setas vermelhas indicam os núcleos celulares corados com hematoxilina, em tons de roxo ou azul, e as setas azuis indicam o epitélio, corado com eosina, em tons de rosa.

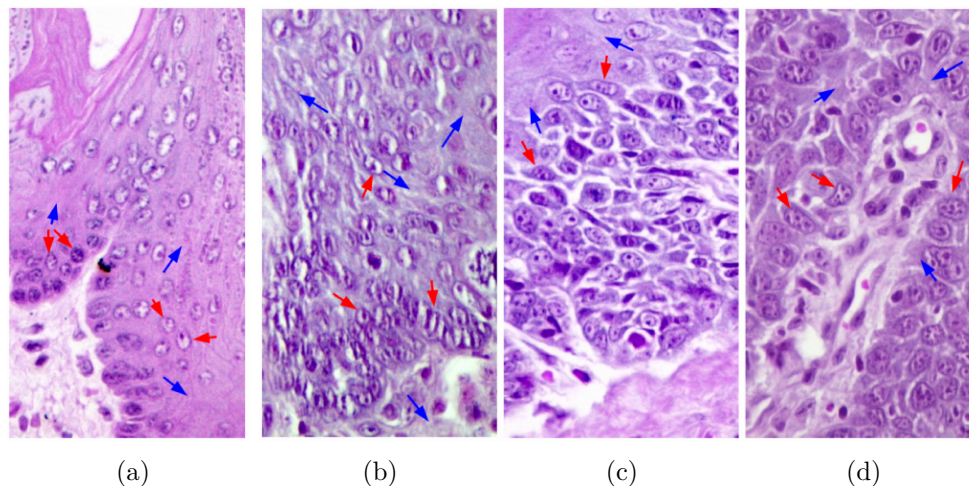


Figura 1 – Exemplos de tecidos histológicos orais nos diferentes níveis de displasia: (a) Saudável, (b) Leve, (c) Moderada e (d) Severa. As setas vermelhas indicam os núcleos celulares e as azuis indicam o epitélio (Figura extraída de (A. B. Silva 2019)).

A análise histológica continua sendo uma das ferramentas mais eficazes para a caracterização da severidade das DEOs (Fonseca-Silva et al. 2016). No entanto, a subjetividade presente na avaliação dessas lesões é um dos principais desafios enfrentados na prática clínica. Características como a extensão do epitélio comprometido e a intensidade das alterações nucleares são critérios fundamentais para a classificação, mas podem ocorrer de maneira independente. Isso gera variações diagnósticas, onde duas lesões com comprometimento semelhante da espessura epitelial podem apresentar alterações nucleares diferentes, o que leva a conclusões distintas entre profissionais (Warnakulasuriya et al. 2008; Smith et al. 2009). Como resultado, a taxa de discordância entre avaliadores pode variar de 19% a 38% (MÜLLER, 2018), e mesmo estratégias como o parecer conjunto entre patologistas, embora eficazes para reduzir a variabilidade, demandam mais tempo e aumentam os custos do processo (Speight et al. 2015; Müller 2018).

A detecção precoce e a correta caracterização das DEOs são essenciais para melhorar o prognóstico dos pacientes, uma vez que essas lesões podem evoluir para carcinoma de células escamosas, cuja taxa de mortalidade é em torno de 50% (V. Kumar et al. 2010; M. d. O. Santos et al. 2023). O risco de transformação maligna das DEOs varia de 6% a 36%, sendo influenciado por fatores externos como tabagismo e consumo de álcool (Smith et al. 2009). O uso de tabaco está fortemente relacionado ao desenvolvimento dessas lesões, com maior incidência de alterações displásicas em indivíduos fumantes em comparação a não fumantes (Santos Pereira et al. 2011; Dost et al. 2014). Esse risco é

potencializado quando o tabaco é combinado ao álcool, podendo atingir até 65% de chance de malignização em casos de DEOs moderadas ou severas (Smith et al. 2009; M. d. O. Santos et al. 2023).

Certas regiões anatômicas são mais vulneráveis à formação dessas lesões, sendo o assoalho bucal um dos locais mais suscetíveis. Isso se deve à fina espessura do epitélio nessa região e à tendência de acúmulo de substâncias nocivas provenientes do álcool e do tabaco, dissolvidas na saliva (Kademani 2019). Diante desse cenário, torna-se evidente a importância de critérios objetivos e técnicas auxiliares para apoiar a tomada de decisão clínica e reduzir a variabilidade diagnóstica entre especialistas.

2.2 Imagens Histológicas com Corantes H&E

A histologia envolve a investigação de tecidos biológicos, abrangendo tecidos humanos, animais e vegetais, com o objetivo de compreender como esses tecidos se organizam para a formação de órgãos. Essa área de estudo concentra-se na análise das células que constituem os tecidos, órgãos e sistemas orgânicos. Devido ao tamanho diminuto das células, é necessário o uso de microscópios para examiná-las em seu estado natural. Para realizar a análise e observações microscópicas dos tecidos, empregam-se etapas que envolvem a sua preparação, seguindo uma série de procedimentos, incluindo a coloração (Vahadane et al. 2016).

As principais etapas para análise de imagens histológicas são representadas na Figura 2. A preparação da amostra histológica começa com a extração do tecido pela biópsia. Após esse processo, a amostra deve passar pela etapa de fixação, onde os tecidos são imersos em soluções químicas para retardar a deterioração e preservar sua estrutura e composição molecular. A escolha dessas soluções depende do tipo de tecido e da análise que será realizada (D. F. D. d. Santos 2022).

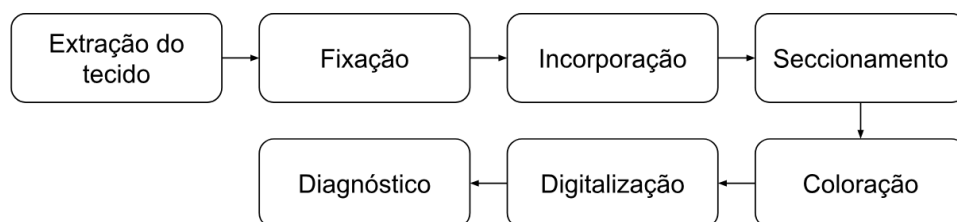


Figura 2 – Principais etapas para análise de imagens histológicas (Figura adaptada de (D. F. D. d. Santos 2022)).

Em seguida, a amostra é incorporada a outras substâncias químicas para eliminar água e outros produtos usados na etapa anterior, o que poderá facilitar a observação das estruturas histológicas pelo microscópio. A etapa de seccionamento envolve a impregnação do tecido com parafina, permitindo a realização dos cortes em secções finas e rígidas. Após

o corte, essas secções são colocadas em lâminas de vidro e submetidas à coloração (TOSTA et al. 2019b).

A coloração dos tecidos é realizada utilizando corantes que diferenciam os componentes básicos e ácidos das células. Corantes comumente utilizados são baseados em H&E. A eosina cora o citoplasma e o colágeno em tons de rosa, enquanto a hematoxilina cora os núcleos celulares e outros componentes ácidos em tons de roxo ou azul. Após a coloração, uma lâmina de vidro é colocada sobre a amostra do tecido para protegê-la. Por fim, a amostra pode ser digitalizada, utilizando uma câmera acoplada a um microscópio ou por digitalizadores (A. B. Silva 2019).

Um problema comum na análise de tecidos é a variação indesejável na cor dessas amostras devido a diferenças na representação por diferentes digitalizadores, matérias-primas, além de técnicas de fabricação de fornecedores de corantes e protocolos de preparação empregados em diferentes laboratórios de patologia (Vahadane et al. 2016). Isso é um obstáculo, pois, desse modo, a coloração de cada amostra poderá ser diferente, o que pode ter consequências nas avaliações de especialistas, tornando-as não unânimes. Além disso, o nível de experiência e carga de trabalho do patologista pode influenciar a qualidade da análise das regiões da imagem (Chan et al. 2020b). Nos últimos anos, tem-se observado o uso de algoritmos computacionais de processamento de imagens e visão computacional para auxiliar patologistas no processo de análise das amostras de tecido digitalizadas (A. B. Silva et al. 2024a).

2.3 Sistemas de Apoio a Diagnósticos

Sistemas CAD são ferramentas especializadas que têm como objetivo auxiliar especialistas no processo de diagnóstico de pacientes. Esses sistemas oferecem suporte ao fornecer informações adicionais que contribuem para a tomada de decisão final do diagnóstico, fornecida pelo especialista (Yeasmin et al. 2024). Essa análise computacional pode ser útil, uma vez que a decisão de patologistas pode variar devido a fatores intra e interpessoais (Chan et al. 2020b).

Os sistemas CAD possuem a aplicação principal de auxiliar na detecção de lesões por meio da identificação de padrões em imagens médicas, auxiliando no diagnóstico ao realizar a classificação de características identificadas nas imagens em padrões normais ou anormais. Esses sistemas podem utilizar técnicas provenientes de diferentes áreas de pesquisa, como visão computacional, que abrange a segmentação e a extração de características das imagens, e inteligência artificial, que se dedica a métodos capazes de reconhecer padrões e realizar classificações (Yeasmin et al. 2024).

Os escopos de aplicação dos sistemas CAD contemplam, dentre outros, a área da mamografia e o diagnóstico de tecidos epiteliais, porém eles também podem ser aplicados a diversas outras modalidades de obtenção de imagens, como radiografia, tomografia,

ressonância magnética, ultrassonografia, medicina nuclear e histologia (Junior 2018).

Na Figura 3, são ilustradas as etapas de um sistema CAD, o qual pode ser dividido em cinco etapas distintas: aquisição das imagens, onde as imagens histológicas são obtidas seguindo o protocolo apresentado na seção anterior; pré-processamento, onde diversas metodologias de melhoria de imagem podem ser aplicadas, como a normalização de cores dos corantes; segmentação, que tem o intuito de identificar as regiões de interesse de uma determinada imagem; extração de características, onde as características interessantes para um determinado problema de classificação são extraídas; e, por fim, a classificação, que utiliza as informações obtidas pelas etapas anteriores para avaliar as imagens analisadas e, no contexto de histologia, realizar o diagnóstico de um determinado tecido (W. He et al. 2022).

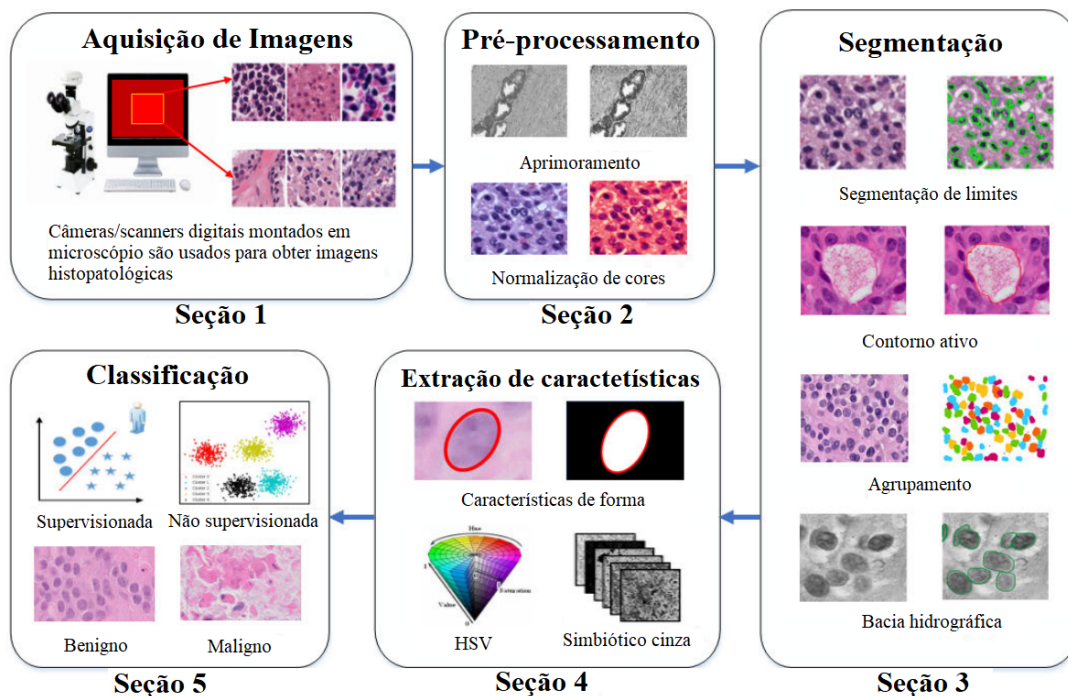


Figura 3 – Etapas e exemplos de técnicas realizadas em um sistema CAD (Figura adaptada de (W. He et al. 2022)).

Os sistemas CAD dependem do processo de aprimoramento da imagem, na etapa de pré-processamento, para prepará-la para uma análise mais aprofundada em etapas posteriores (Ramadan 2020). Em geral, a fim de realizar o aprimoramento da imagem, são utilizados algoritmos para normalização de cores, filtros para remoção de ruído e operações de borrimento com a finalidade de realçar as estruturas histológicas (A. B. Silva 2019). Os métodos de normalização modificam as cores da imagem analisada usando as características cromáticas de uma imagem de referência específica. A normalização já demonstrou sua efetividade ao permitir o aumento do desempenho de sistemas em patologia digital quando aplicada (C.-M. Chen et al. 2020).

A segmentação de imagens opera a nível de pixels, dividindo a imagem em áreas

distintas com propriedades únicas e delimitando regiões de interesse, o que é essencial para melhorar a quantificação e análise das imagens (Yeasmin et al. 2024; Chan et al. 2020b). Investigar se erros na normalização podem afetar a segmentação ainda é uma questão em aberto. A presença de ruído em uma imagem gera incerteza, dificultando seu processo de classificação. Isso ocorre porque o ruído altera os valores de intensidade dos pixels, prejudicando a uniformidade das faixas de intensidade da imagem. Se essas alterações impactam diretamente na segmentação e análise da imagem é um ponto ainda a ser explorado (Xintong Li et al. 2022).

2.4 Normalização de Imagens com Corantes H&E

Os sistemas CAD desempenham um papel fundamental nas etapas de análise e processamento de imagens histológicas, incluindo pré-processamento, segmentação, extração de características e classificação. O processo de normalização de cores ocorre no pré-processamento, juntamente com outros aprimoramentos das imagens, como redução de ruídos, manipulação de dados ausentes (para solucionar problemas de *missing data*) e aumento do contraste (Ramadan 2020).

A normalização de cor, no contexto de processamento de imagens histológicas, tem como objetivo ajustar os canais de cores, oferecendo um alto contraste entre as estruturas representadas e preservando as características e informações da imagem original (Salehi et al. 2020). O resultado da normalização é a obtenção de imagens com coloração semelhante entre si. Isso permite que os sistemas classificadores compreendam de forma mais precisa as diferentes estruturas celulares e teciduais no contexto das imagens histológicas. Na Figura 4, é ilustrado um exemplo de normalização de imagens histológicas, sendo possível realizar esse processamento a partir de uma imagem de referência (a) e uma imagem original (b). Nesse exemplo, nota-se que a imagem resultante (c) adquire a coloração da imagem de referência, sofrendo alteração em seus canais de cores, mas mantendo as estruturas celulares presentes na imagem original.

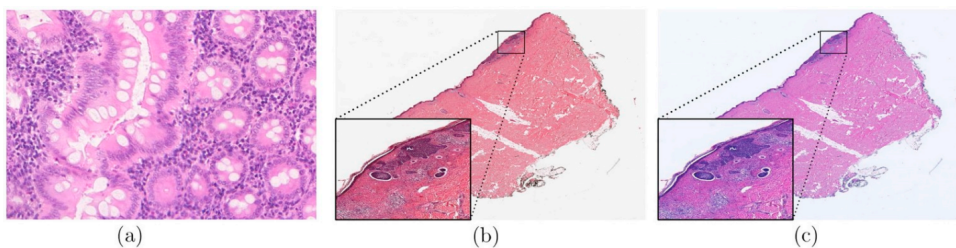


Figura 4 – Exemplo de normalização de uma imagem original (b), dada uma imagem de referência (a), originando uma imagem normalizada (c) (Figura extraída de (TOSTA et al. 2019a)).

Os desafios da área de normalização de imagens incluem (i) preservação de estruturas, pois as estruturas da imagem original devem ser mantidas após a normalização (Vahadane

et al. 2016), (ii) aprimoramento de contraste, auxiliando a identificação de estruturas e diferenças entre os corantes utilizados (Lei et al. 2020), e (iii) evitar a adição de ruídos, que podem alterar as estruturas representadas e prejudicar o desempenho de etapas de processamento posteriores (Pérez-Bueno et al. 2022).

A aplicação da normalização entre os corantes H&E pode considerar algumas importantes propriedades biológicas em sua definição metodológica. É essencial assumir que a densidade de um corante e a densidade óptica associada a ele podem estar ausentes (zero) ou presentes (valores positivos) em uma determinada estrutura histológica. Porém, essas densidades não podem ser negativas, pois a existência de uma densidade de coloração negativa implica na emissão de luz, condição biologicamente inválida. Além disso, uma estrutura biológica, como um núcleo, pode se relacionar a mais de um corante químico, como no caso da coloração H&E, indicando assim a importância de considerar a esparsidade, ou mistura, dos corantes no processo de normalização das imagens (Vahadane et al. 2016).

Algumas técnicas de normalização utilizam uma única imagem de referência selecionada por especialistas, considerada um exemplo adequado de coloração dos corantes. Também é possível realizar a normalização sem o uso da imagem de referência, com ajustes de cor apenas pelo processamento das imagens originais, ou por conjuntos de treinamento em técnicas de aprendizagem profunda. Assim, esses métodos de normalização podem ser categorizados como correspondência de histogramas, transferência de cores e correspondência espectral (TOSTA et al. 2019b).

Os algoritmos de correspondência de histograma são baseados na correspondência dessas representações no espaço de cores *Red, Green and Blue* (RGB), sendo realizada em cada um dos canais vermelho, verde e azul, respectivamente. Os métodos baseados em correspondência de histograma podem levar a colorações errôneas de regiões histológicas, o que leva a uma preservação inadequada das informações histológicas da imagem original pela normalização (Xingyu Li et al. 2015).

Na classe de transferência de cores, são computadas estatísticas de cada canal de cor entre uma imagem original e uma imagem de referência. No entanto, essa abordagem aplicada separadamente aos canais possui duas limitações. Primeiro, as fontes de variação de cor nas imagens de histopatologia são agrupadas e não tratadas separadamente. Segundo, as estatísticas intrinsecamente relacionadas às informações histológicas nas imagens são modificadas, o que pode resultar na perda das características do tecido após a normalização (Xingyu Li et al. 2015).

Os métodos de correspondência espectral baseiam-se em estimativas da representação espectral de corantes em cada canal de cor e na concentração desses corantes nos pixels da imagem. É importante que essas estimativas sejam robustas e precisas, a fim de preservar as estruturas histológicas e evitar a introdução de ruído na imagem normalizada (Xingyu Li et al. 2015). A correspondência espectral possui vantagens em relação a outras técnicas

devido ao seu potencial de preservar as características teciduais, um critério essencial na avaliação da normalização (TOSTA et al. 2019a). No contexto de imagens coradas com H&E e da correspondência espectral de imagens, pode-se abordar o desafio da normalização de cores como um problema de aprendizado de dicionários, de modo que o dicionário estimado represente a coloração da imagem.

2.5 Métodos de Aprendizado de Dicionários

Modelos para representação das propriedades dos corantes podem ser definidos por metodologias baseadas em aprendizado de dicionários (Macenko et al. 2009). O objetivo do aprendizado de dicionário é obter uma representação esparsa da entrada, representando-a como uma combinação linear de elementos-chave que compõem um dicionário. Esses elementos, também conhecidos como átomos, são considerados mais flexíveis em comparação com uma base ortogonal, pois oferecem representações mais ricas dos dados (Dhaini et al. 2023).

Um dicionário, no contexto de normalização, é uma base aprendida de padrões cromáticos e texturais usada como referência para reconstituir imagens em um padrão visual consistente, promovendo a normalização de cores de forma controlada e replicável. Na literatura, são propostos diferentes métodos para aprendizado de dicionários que são adequados para representar classes específicas de sinais e dados provenientes de múltiplos sensores. Esses dicionários são projetados com o intuito de capturar as características mais relevantes dos dados, permitindo uma representação eficiente e eficaz dos mesmos (Tosic et al. 2011).

Um exemplo de aprendizado de dicionário é a técnica *Non-negative Matrix Factorization* (NMF), proposta, no contexto de processamento de imagens histológicas, por (Rabinovich et al. 2003). Sem o auxílio de um algoritmo, é necessário que um especialista indique manualmente pontos na imagem do tecido que representem espectros de corantes puros. Este procedimento requer intervenção humana e, portanto, é demorado e sujeito a erros (Rabinovich et al. 2003). Os autores então apresentaram o primeiro sistema capaz de realizar essa decomposição de cores de forma automatizada, pela NMF.

Em cenários ideais, o perfil espectral de cada corante deve estar contido em uma das bandas espectrais, com absorções diretamente representativas das concentrações do corante. Realisticamente, no entanto, o perfil espectral dos corantes se sobrepõe e se estende por várias bandas espectrais, e o objetivo de recuperar os componentes que representam as porcentagens do corante requer uma análise mais cuidadosa. A decomposição de cores dos corantes pode ser formulada como um problema de fatoração de matrizes conforme:

$$V = WH, \tag{1}$$

em que V é a matriz de densidades ópticas e representa uma imagem nos canais RGB, portanto de dimensões $3 \times N$, onde N representa o conjunto de pixels da imagem. W é a matriz de aparência de cores dos corantes, de dimensão 3×2 no caso de imagens H&E, com representação de cada corante por cores RGB. A variável H é o mapa de densidade de corantes, de dimensão $2 \times N$, isto é, uma matriz que relaciona os corantes H&E com o conjunto de pixels da imagem, demonstrando a concentração de cada corante para cada pixel da imagem.

A fim de realizar a normalização de uma imagem, estima-se a matriz W de uma imagem de referência com boa coloração e proporção de corantes. Torna-se então possível combiná-la com o mapa de densidade de corantes de uma imagem original pela Equação 1, resultando em uma matriz de densidades ópticas normalizada. Após converter essa matriz de densidades para o espaço de cores RGB, o resultado é a imagem normalizada, mantendo as estruturas da imagem original, mas com a aparência de cores dos corantes da imagem de referência.

Os autores de (Vahadane et al. 2016) introduziram o termo esparsidade à normalização de cores de imagens histológicas, supondo que um tipo de corante está vinculado às estruturas histológicas em diferentes proporções. Assume-se que o tecido representado pelos pixels da imagem é caracterizado por sua coloração efetiva, permitindo que cada pixel seja modelado como uma mistura esparsa dos corantes. Uma estrutura biológica, como um núcleo, pode se relacionar com mais de uma coloração química, como no caso da coloração H&E, mas suas proporções relativas são fixadas para uma determinada estrutura biológica dentro de uma imagem.

Denomina-se essa mistura de corantes que está exclusivamente presente em um tipo de estrutura biológica como um corante eficaz. Assim, estimam-se os vetores de base de coloração efetivos e não os vetores de base de coloração pura para definição da matriz W . Essa abordagem é denominada Fatoração de Matriz Não-negativa Esparsa (do inglês, *Sparse Non-negative Matrix Factorization* (SNMF)). A Figura 5 representa graficamente vetores de coloração efetivos, ilustrando o núcleo celular corado com H efetivo, que é uma mistura de uma grande quantidade de H puro e uma pequena quantidade de E puro, e em contraste, o tecido conjuntivo corado com E efetivo é uma mistura de uma grande quantidade de E puro e uma pequena quantidade de H puro.

2.6 Redes Neurais Convolucionais

Redes neurais convolucionais (do inglês, *Convolutional Neural Networks* (CNN)) são uma forma de realizar processamento e análise de imagens por diferentes modelos. Para compreender as CNNs, é necessário conhecer também o conceito de redes neurais artificiais (do inglês, *Artificial Neural Networks* (ANN)). De forma análoga à rede de neurônios presente no cérebro humano, as ANNs são capazes de receber dados na camada de entrada,

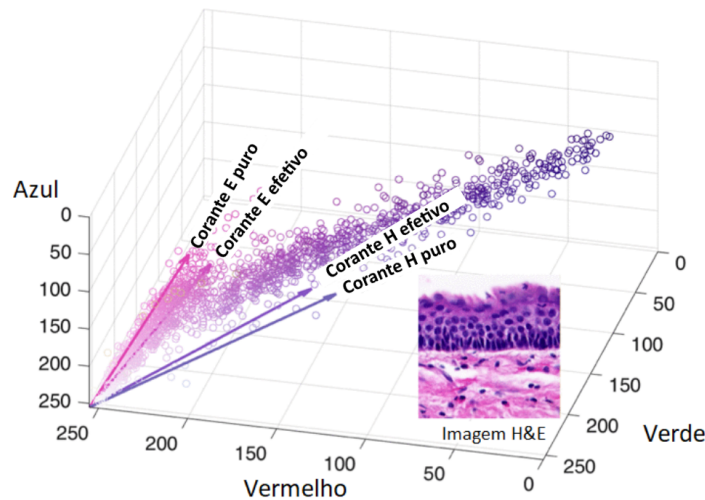


Figura 5 – Representação gráfica dos vetores de coloração efetivos (Figura adaptada de (Vahadane et al. 2016)).

processá-los e transmiti-los para as camadas subsequentes até a camada de saída. Essa comunicação entre as camadas permite sua aplicação em uma variedade de problemas de inteligência artificial (Goodfellow et al. 2016).

Os neurônios artificiais são as estruturas básicas de uma ANN, constituídos por três componentes principais: (i) sinapses, correspondentes a conexões que desempenham o papel de transmitir informações entre os neurônios, (ii) uma função responsável por combinar as informações recebidas pelas sinapses e gerar uma nova informação, e (iii) uma função de ativação, que transforma a informação produzida pelo neurônio em um valor finito, permitindo determinar a intensidade do estímulo propagado (Haykin 2009; A. B. Silva 2019).

Na Figura 6, pode-se observar uma ANN simples onde, no contexto de histologia, as imagens dos tecidos são recebidas na camada de entrada e os neurônios nas camadas ocultas ou intermediárias são responsáveis por calcular os pesos das informações recebidas através de uma função de ativação. Diversas funções de ativação podem ser utilizadas, incluindo a função linear, a função sigmoide e a Unidade Linear Retificada (do inglês, *Rectified Linear Unit* (ReLU)) (A. B. Silva 2019).

Os pesos são transmitidos pelos neurônios até a camada de saída, onde são parametrizados em uma função de classificação. Uma das funções geralmente utilizadas é o classificador *softmax*, sendo uma generalização da forma binária da regressão logística definida para mapear um conjunto de dados de entrada para rótulos de classe de saída por meio do produto escalar dos dados e da matriz de pesos (Rosebrock 2017).

As CNNs são um tipo de ANN desenvolvidas especialmente para processamento de imagens e visão computacional. A principal distinção entre as CNNs e as ANNs convencionais é a presença das camadas convolucionais. Geralmente, uma CNN é composta por camadas convolucionais, camadas de *pooling* e camadas totalmente conectadas (do

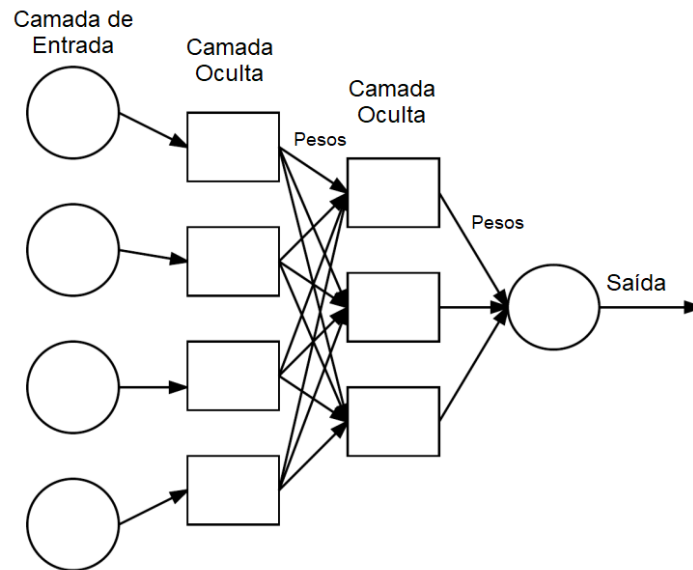


Figura 6 – Uma arquitetura simples de rede neural artificial (Figura adaptada de (Rosebrock 2017)).

inglês, *Fully-Connected* (FC)). Enquanto as camadas convolucionais e de *pooling* são responsáveis pela extração de características, as camadas FC realizam a classificação dessas características, gerando um rótulo de saída para os dados de entrada (Goodfellow et al. 2016).

A camada convolucional utiliza um processo linear que envolve a multiplicação dos dados de entrada por matrizes numéricas conhecidas como *kernels*. Essas matrizes processam diferentes regiões da imagem, onde cada posição da matriz representa um pixel da imagem de entrada, aplicando um filtro por meio da operação de convolução. Essa operação é ilustrada pela Figura 7, onde um *kernel* de tamanho 3×3 pixels, representado em vermelho, percorre a matriz de pixels da imagem, representada em branco (Rosebrock 2017).

131	162	232	84	91	207
104	-1	10	+1	237	109
243	-2	20	+2	135	26
185	-15	20	+1	61	225
157	124	25	14	102	108
5	155	16	218	232	249

Figura 7 – Funcionamento de um *kernel* sobre uma região da imagem (Figura extraída de (Rosebrock 2017)).

Diferentes valores dos *kernels* representam diferentes filtros (Rosebrock 2017). A Figura 8 exemplifica o resultado sobre diferentes transformações em uma imagem a partir de convoluções. Com essas diferentes versões da imagem original, a rede é capaz de extrair características e processá-las a fim de realizar uma classificação da imagem, por exemplo.

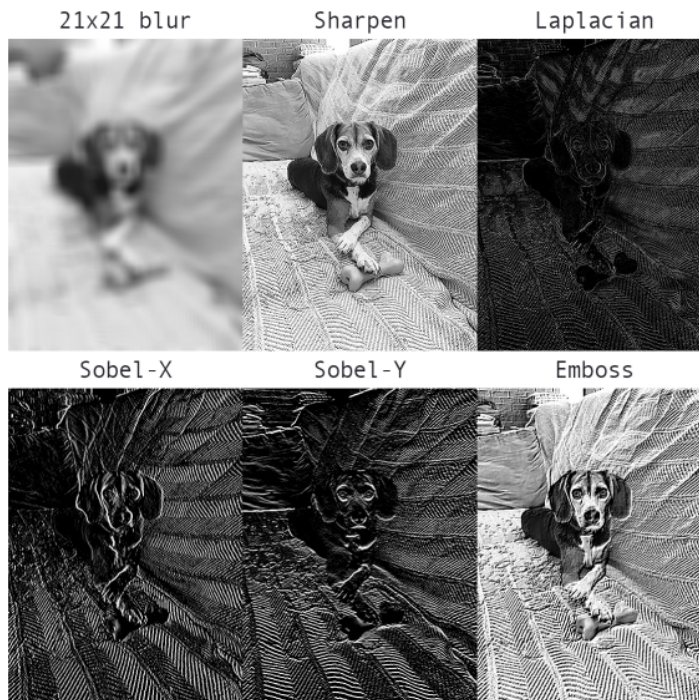


Figura 8 – Exemplos de transformações por diferentes tipos de *kernels* (Figura extraída de (Rosebrock 2017)).

Após as convoluções serem aplicadas nas imagens, é gerada uma grande quantidade de valores, que são repassados para as funções de ativação nas camadas FC. O processamento de um número significativo de dados exige um alto esforço computacional. Assim, as camadas de *pooling* foram desenvolvidas para diminuir esse esforço, reduzindo as dimensões das imagens, mas mantendo suas características mais importantes (Goodfellow et al. 2016). Para isso, é utilizado um algoritmo semelhante à camada convolucional: uma janela pequena de tamanho variável, chamada *pool*, que desliza sobre a matriz de entrada, mas não se aplica uma convolução. Em vez disso, aplica-se uma função de redução. As principais funções de redução são a *max pooling* e *average pooling* (D. F. D. d. Santos 2022). A Figura 9 exemplifica a utilização dessas funções de redução, onde o *max pooling* seleciona o maior valor da *pool*, e o *average pooling* calcula a média entre os valores da *pool* (Rosebrock 2017).

Em geral, as primeiras camadas das CNNs são responsáveis por extrair as características da imagem original, enquanto as camadas finais consistem em neurônios também utilizados em ANNs convencionais. As camadas FC consistem em uma série de camadas em que todas as saídas estão conectadas às entradas da próxima camada por meio de pesos adaptados durante o treinamento, geralmente utilizando o algoritmo de retropropagação.



Figura 9 – Exemplos das operações *max pooling* (b) e *average pooling* (c) em uma matriz de tamanho 4×4 pixels (a), utilizando *pool* de tamanho 2×2 pixels (Figura extraída de (D. F. D. d. Santos 2022)).

Nas tarefas de classificação, as saídas correspondem às probabilidades de cada classe, obtidas por meio de uma função de ativação, sendo a função *softmax* a mais comumente utilizada para essa finalidade (Rosebrock 2017).

Existem diversos modelos de CNNs disponíveis, e a literatura apresenta diferentes formas de combinar essas camadas, oferecendo uma ampla gama de modelos de CNNs para pesquisas da área (D. F. D. d. Santos 2022). Exemplos de modelos de CNNs incluem o *LeNet*, *AlexNet*, *VGG16* e *ResNet*, entre outros. A Figura 10 ilustra a arquitetura *LeNet*, que foi o primeiro modelo de CNN que obteve sucesso na tarefa de reconhecimento de caracteres escritos à mão, e a Figura 11 ilustra a arquitetura *AlexNet*, que foi o responsável por popularizar o uso de modelos CNN no campo de visão computacional.

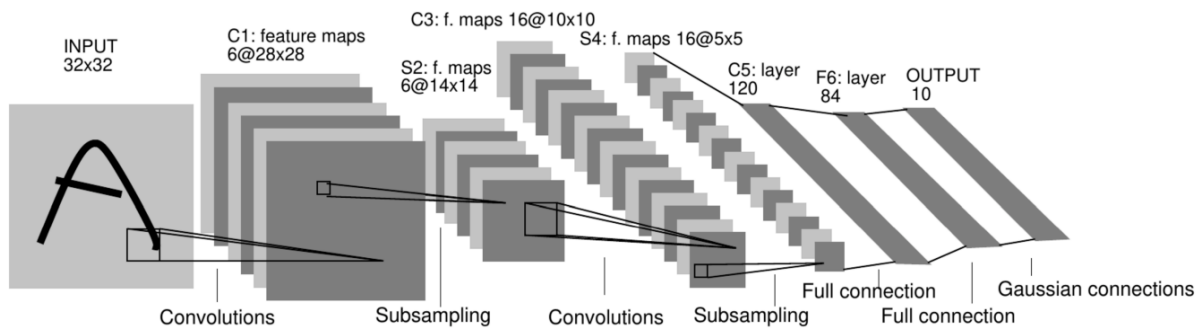


Figura 10 – Arquitetura do modelo *LeNet*. (Figura extraída de (Lecun et al. 1998)).

A rede *LeNet* recebe imagens de tamanho 32×32 pixels, possui duas camadas convolucionais, duas camadas de *pooling* ou *subsampling* e duas camadas totalmente conectadas, gerando uma saída de dez classes (Lecun et al. 1998). Já a *AlexNet* é um modelo mais robusto, recebendo imagens de tamanho 224×224 pixels, com cinco camadas convolucionais, três camadas de *pooling* e duas camadas totalmente conectadas, gerando uma saída de mil classes (Krizhevsky et al. 2017).

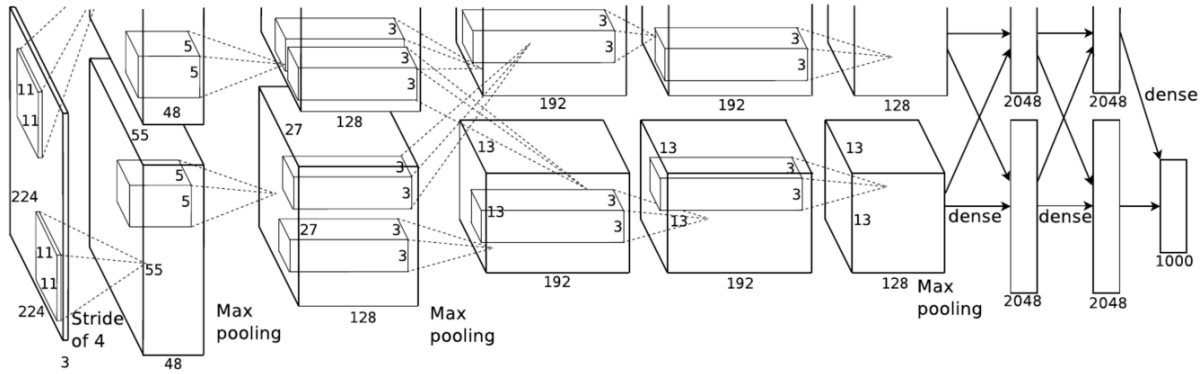


Figura 11 – Arquitetura do modelo *AlexNet*. (Figura extraída de (Krizhevsky et al. 2017)).

2.7 Segmentação de Imagens

A segmentação de imagens é a área de pesquisas responsável por identificar e separar objetos diferentes em uma imagem, utilizando variadas metodologias. A segmentação é caracterizada como um tipo de tarefa da Visão Computacional (ou do inglês, *Computer Vision*) (Hao et al. 2020). Existem cinco tarefas principais de visão computacional, cada uma possuindo seus objetivos distintos. São elas: segmentação semântica, classificação de imagens, detecção de objetos, segmentação de instância e segmentação panóptica (Zhao et al. 2019). A Figura 12 exemplifica cada uma dessas principais tarefas.

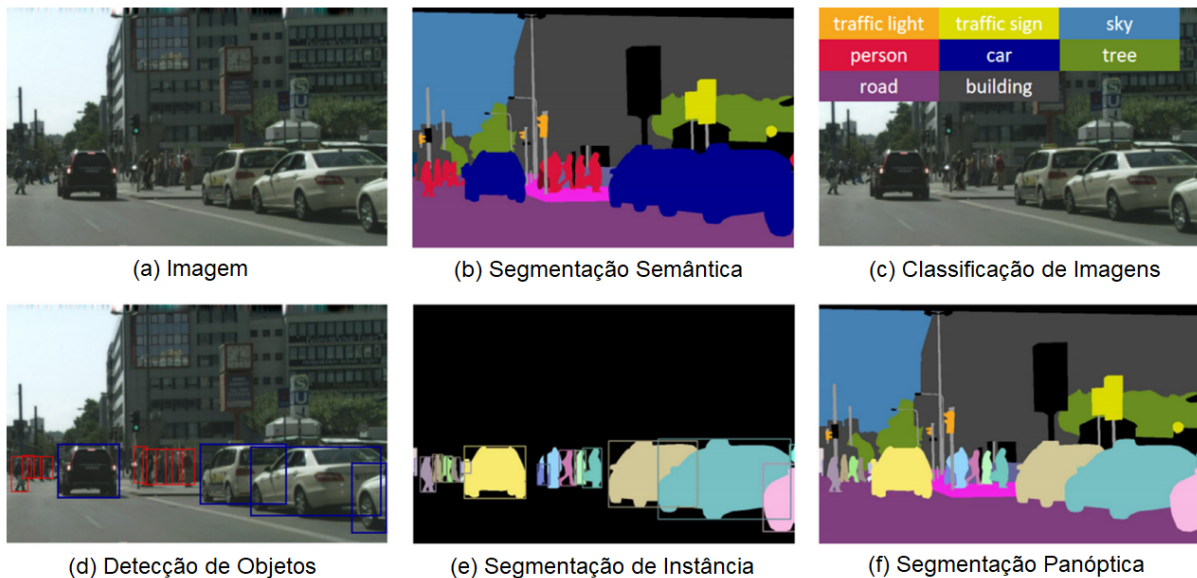


Figura 12 – Exemplo das principais tarefas de visão computacional (Figura adaptada de (Hao et al. 2020)).

A segmentação semântica visa particionar uma imagem em subconjuntos mutuamente exclusivos, nos quais cada subconjunto representa uma região significativa da imagem original. Assim, ela tem o objetivo de atribuir um rótulo de categoria predefinido a cada pixel da imagem, segmentando os elementos de uma imagem em duas ou mais classes

(Minaee et al. 2021).

A classificação de imagens tem o objetivo de identificar os objetos presentes em uma imagem, sem o objetivo de identificar exatamente onde esses objetos se encontram. Ela visa atribuir um ou mais rótulos de categoria para uma imagem inteira. Em outras palavras, um algoritmo de classificação de imagens nos diz quais objetos existem em uma imagem (Zhao et al. 2019).

A detecção de objetos vai além da tarefa de classificação, e identifica os objetos das classes em retângulos na própria imagem. Desse modo, além de identificar quais objetos estão presentes na imagem, também é possível definir a posição em que o objeto se encontra na imagem (Zhao et al. 2019).

A segmentação de instância visa detectar cada objeto como um indivíduo na imagem, isto é, por mais que o objeto pertença a uma classe, ele é único em relação aos outros objetos da mesma classe. Na Figura 12, isso é identificado pelas cores atribuídas aos objetos e aos seus respectivos retângulos, onde cada carro e pessoa possui uma cor única, por mais que pertençam a uma determinada classe (Minaee et al. 2021).

Por fim, a detecção panóptica tem o objetivo mais complexo, que atribui um rótulo semântico e um rótulo de instância a cada pixel. Comparadas com os requisitos tradicionais de segmentação, essa tarefa é mais desafiadora, pois seu alvo se estende de objetos semânticos e instâncias de objetos para uma combinação dos dois (Hao et al. 2020).

No contexto da segmentação de imagens histológicas, um dos problemas abordados é a segmentação dos núcleos celulares dos tecidos observados, de modo a separá-los do restante da imagem. Para isso, o método de segmentação semântica é o mais apropriado, visto que é necessário identificar uma única classe de objeto, sendo este o núcleo celular (Nasir et al. 2023). Neste trabalho, o foco de estudo foi essa tarefa e como mudanças nas imagens podem alterar os resultados das segmentações semânticas.

Para realizar a segmentação semântica de imagens, pode-se utilizar redes com arquitetura *encoder-decoder*, e variações da rede neural convolucional U-Net são desenvolvidas e utilizadas com esse intuito (Nasir et al. 2023; Basu et al. 2024). A rede tem esse nome devido ao padrão “U” simbolizado pelo padrão de suas camadas. Sua representação é ilustrada na Figura 13

A arquitetura da U-Net é dividida em duas metades espelhadas. A primeira metade, também chamada de *encoder*, realiza convoluções convencionais e *max pooling*, diminuindo o tamanho da amostra. Já a segunda parte, chamada de *decoder*, realiza operações de *upsampling* (ou aumento de amostragem), reconstruindo a resolução da imagem de entrada enquanto preserva as informações das características aprendidas no *encoder*. As camadas espelhadas do *encoder* e *decoder* são combinadas por uma operação denominada *skip connection*, que concatena o resultado dessas camadas. Desse modo, as informações detalhadas das camadas iniciais do *encoder* são transferidas diretamente para as camadas correspondentes no *decoder*, fazendo então com que a rede seja capaz de aprender a

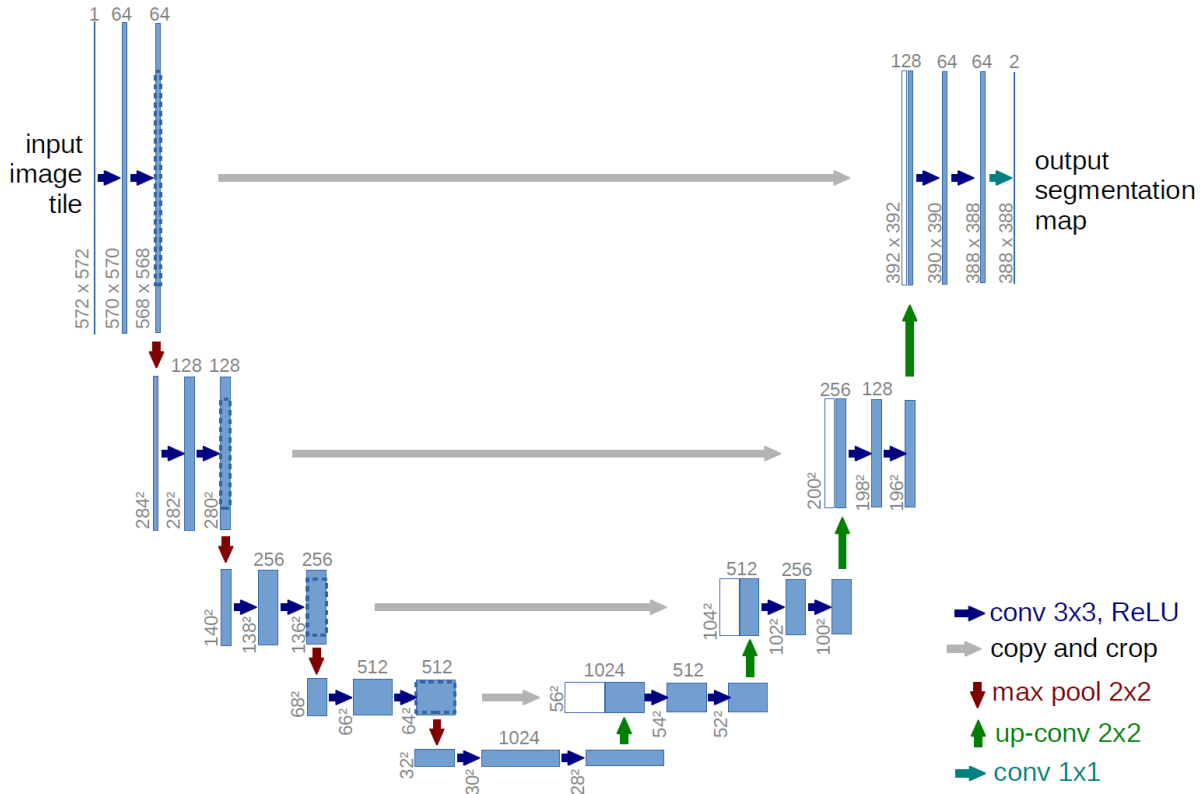


Figura 13 – Exemplo de arquitetura de uma rede U-Net (Figura extraída de (Ronneberger et al. 2015)).

montar uma saída mais precisa com base nessas informações (Ronneberger et al. 2015; Asgari Taghanaki et al. 2021).

2.8 Explicabilidade de Redes Neurais Artificiais

A popularização de modelos complexos de aprendizado de máquina, como redes neurais profundas, trouxe ganhos significativos em desempenho preditivo, mas também evidenciou uma limitação crucial: a dificuldade de compreender como esses modelos tomam decisões. Em muitas aplicações, especialmente nas que envolvem alto impacto social, como saúde, finanças ou justiça, essa compreensão é essencial. Nesse contexto, surge a área de explicabilidade em aprendizado de máquina, cujo objetivo é tornar os modelos mais transparentes, compreensíveis e confiáveis para usuários humanos (Minh et al. 2022).

A interpretabilidade dos modelos de aprendizado de máquina não é apenas desejável, mas essencial para possibilitar auditoria, depuração e melhoria contínua dos sistemas. Mesmo em aplicações de baixo risco, como sistemas de recomendação de filmes, a capacidade de compreender o comportamento do modelo é valiosa tanto na fase de pesquisa quanto após sua implantação. Quando um modelo entra em produção, erros inevitavelmente surgem, e a possibilidade de explicá-los ajuda a identificar suas causas e a orientar correções. Um exemplo é o de um classificador treinado para distinguir lobos de *huskies*,

que passou a associar erroneamente a presença de neve nas imagens como um indicativo de “lobo”. Embora essa associação fizesse sentido no conjunto de treinamento, ela falha no uso real, revelando um problema de generalização. Ao tornar as decisões dos modelos explicáveis, torna-se mais fácil avaliar aspectos cruciais como justiça, privacidade, robustez, causalidade e confiança. Isso permite identificar se o modelo está reproduzindo vieses, expondo dados sensíveis, sendo sensível a pequenas variações na entrada, aprendendo relações incorretas ou dificultando a confiança dos usuários por atuar como uma caixa-preta (Molnar 2022).

Embora os termos explicabilidade e interpretabilidade sejam frequentemente usados como sinônimos, há uma distinção sutil, porém importante, entre eles. A interpretabilidade refere-se à capacidade de um modelo ser entendido diretamente. Por exemplo, um modelo linear ou uma árvore de decisão simples pode ser interpretado observando-se seus coeficientes ou regras. Já a explicabilidade trata da geração de explicações sobre o funcionamento ou decisões de modelos que, por si só, não são interpretáveis. Em outras palavras, explicabilidade envolve métodos e técnicas que auxiliam na compreensão de modelos que funcionam como verdadeiras caixas-pretas (Molnar 2022).

Redes neurais são um exemplo clássico de modelos de difícil interpretação. Devido à sua natureza altamente parametrizada, à presença de múltiplas camadas não lineares e à forma como aprendem representações complexas, esses modelos não oferecem caminhos claros para a análise humana. Apesar disso, é possível e necessário buscar formas de extrair explicações sobre o comportamento da rede, seja para entender uma decisão individual, seja para avaliar o modelo como um todo. A busca por explicabilidade em redes neurais representa uma resposta direta à necessidade de tornar modelos complexos mais transparentes e auditáveis (Minh et al. 2022). No entanto, essa tarefa não é trivial.

Existem desafios técnicos e conceituais importantes para a explicabilidade, como a confiabilidade das explicações, sua consistência entre diferentes métodos e sua utilidade para tomadores de decisão humanos. Além disso, há um equilíbrio delicado entre a complexidade de um modelo e o grau de explicabilidade que se consegue extrair dele. Por esse motivo, a explicabilidade não deve ser vista apenas como uma ferramenta adicional, mas como parte essencial do processo de desenvolvimento e validação de sistemas baseados em aprendizado de máquina (Molnar 2022; Minh et al. 2022).

2.9 Considerações Finais

Este capítulo apresentou os conceitos teóricos relacionados ao tema do estudo e à metodologia proposta. Foram explorados os tópicos que abrangem as etapas e abordagens aplicadas ao trabalho, destacando a necessidade de compreender as características e os desafios específicos para o avanço das próximas fases da pesquisa. Entre os aspectos enfatizados, estão a compreensão do contexto das imagens histológicas coradas com H&E

e dos sistemas CAD, os conceitos de normalização de imagens utilizando métodos baseados em aprendizado de dicionários e a segmentação de imagens histológicas utilizando CNN, por meio do modelo U-Net.

Trabalhos Relacionados: Estado da Arte

Este capítulo apresenta estudos focados na normalização de corantes H&E baseados em métodos de aprendizagem de dicionário e estratégias de segmentação de imagens histológicas de displasia. Para embasar o estudo proposto, foram considerados apenas artigos aplicados a imagens histológicas de cavidade oral na etapa de segmentação.

3.1 Normalização de Corantes em Imagens H&E

A normalização de corantes H&E em imagens histológicas é um tema de pesquisa abordado por diversos trabalhos na literatura (Rong et al. 2023; Runz et al. 2021; Salehi et al. 2020; Shaban et al. 2019). Diferentes técnicas computacionais podem ser utilizadas para a normalização dessas imagens, como descrito em (TOSTA et al. 2019b), a fim de padronizar a aparência das amostras de tecidos digitalizados. Nesse contexto, técnicas de aprendizado de dicionário são relevantes ao considerar propriedades biológicas dos corantes e dos tecidos (Pérez-Bueno et al. 2022). Dessa forma, trabalhos relacionados a técnicas de aprendizado por dicionário para normalização de corantes H&E são apresentados nesta seção.

Em 2015, os autores do estudo (Xingyu Li et al. 2015) propuseram um método que calcula estimativas espectrais e de luminância da imagem com o objetivo de realizar a normalização dos corantes H&E. Esse método usa estatísticas ponderadas por saturação para ajustar cores sem aprendizado de padrões. Para lidar com as variações dos corantes, o método explorou uma decomposição espectral por meio de fatoração de matrizes não-negativas (NMF), a qual divide a imagem nos espectros dos corantes e suas respectivas quantidades presentes em cada pixel da imagem do tecido. Para garantir que essa decomposição seja mais confiável e menos sensível a ruídos, os autores introduziram uma abordagem de inicialização baseada em estatísticas que levam em conta a saturação das cores, destacando as regiões mais ricas em cor como mais representativas. Em seguida, a lei de Beer-Lambert foi aplicada para alinhar os espectros da imagem de entrada com a referência, sem alterar a distribuição dos corantes, assegurando a preservação das infor-

mações histológicas originais. Os autores testaram o método em dois conjuntos públicos diferentes, *NIA dataset* e *UCSB dataset*, sendo o primeiro formado por imagens de linfonos, apresentando uma grande variedade na coloração das imagens, e o segundo formado por imagens de tecidos com câncer de mama. O método apresentou resultados visuais pertinentes no âmbito da normalização de cores, sendo capaz de preservar as informações histológicas após a normalização. Entretanto, os autores não utilizaram métricas consolidadas na literatura para avaliação quantitativa, o que impede sua comparação com trabalhos correlatos.

No estudo em (Vahadane et al. 2016), os autores propuseram o método denominado *Structure-Preserving Color Normalization* (SPCN), que envolveu uma estimativa da matriz de aparência dos corantes e do mapa de densidade dos corantes por meio do método SNMF, considerando a esparsidade dos corantes e bases públicas de imagens histológicas H&E de diferentes tipos de tecidos. A estimativa do mapa de densidade de corantes foi realizada pelo método de regressão de ângulo mínimo (do inglês, *Least Angle Regression* - LARS), e a estimativa da matriz de aparência dos corantes foi realizada pela técnica de desvio de coordenadas de bloco livre de parâmetros com *warm restart*. Em ambas as estimativas, a ferramenta *SParse Modeling Software* (SPAMS) foi utilizada. Posteriormente, o mapa de densidade de corantes da imagem original foi combinado com a matriz de aparência dos corantes de uma imagem de referência. A esparsidade introduzida pelo SNMF foi fundamental nesse processo, pois permitiu a representação da mistura dos corantes, contribuindo para sua deconvolução, com representação de cada corante por uma imagem. Os resultados apresentados mostraram que o método SNMF superou o NMF e a decomposição por valores singulares (do inglês, *Singular Value Decomposition* (SVD)), onde a métrica *Relative Root Mean Squared Error* (RRMSE) melhorou cerca de três vezes em comparação ao NMF para os quatro tipos de tecido comparados (estômago, próstata, cólon e bexiga). Em relação ao SVD, o aumento dessa métrica foi de 20-30% em três tipos de tecido.

Em 2019, os autores em (TOSTA et al. 2019a) apresentaram um trabalho com aprimoramento em relação ao método apresentado por (Vahadane et al. 2016). Essa abordagem foi baseada em três etapas: estimativa de um parâmetro de esparsidade, inicialização da matriz de aparência dos corantes e sua estimativa. Esse estudo apresentou a aplicação da normalização para correção das regiões desbotadas das imagens histológicas, com possíveis correções a amostras de tecido com baixo, médio e alto grau de desbotamento dos corantes, em um banco de amostras histológicas coradas com H&E de melanoma cutâneo escaneadas em diferentes períodos de tempo. Ao comparar os valores obtidos pelas métricas FSIM (0,9706) e distância de Canberra (92,13, 104,90 e 98,39 para os canais RGB das imagens com alto desbotamento) com outros métodos da literatura, os autores identificaram que, embora a correção da cor do tecido pela técnica proposta tenha apresentado bom desempenho, eles não obtiveram uma boa representação das estruturas de fundo das

imagens.

Ainda em 2019, no estudo de (TOSTA et al. 2019c), os autores apresentaram um modelo semelhante, porém foi estimado o parâmetro de esparsidade para imagens com diferentes variações de cores, e sobre imagens de bases públicas de tecidos com linfoma, câncer colorretal e câncer de mama. Uma de suas principais inovações está na forma como a matriz de aparência dos corantes (W) foi estimada. Em vez de manter o mapa de densidade dos corantes (H) fixo, como ocorre em métodos anteriores, esse método realiza uma estimativa iterativa da matriz W e do mapa H para cada mini-lote da imagem, utilizando mínimos quadrados ponderados. Esses mini-lotes correspondem a pequenas frações da imagem (0,1% do tamanho da imagem), permitindo atualizações estáveis e localmente adaptadas de W , com auxílio de um algoritmo de gradientes conjugados. Esse processo resulta em uma matriz de corantes mais representativa para cada imagem, respeitando suas particularidades biológicas e estruturais. Outro destaque do método está na definição automática do parâmetro de esparsidade. Em vez de definir esse valor manualmente, como ocorre tradicionalmente, os autores propuseram uma estimativa baseada em lógica *fuzzy*. Essa técnica avalia os graus de pertencimento dos pixels aos corantes H&E a partir da intensidade óptica, utilizando funções de associação parametrizadas. Isso permite que o modelo reflita melhor a representação de cores dos tecidos, onde os corantes tendem a atuar de forma dominante em regiões específicas. Os autores utilizaram a métrica *Quaternion Structural SIMilarity* (QSSIM) para comparar os resultados com outros estudos, sobressaindo-se sobre eles com os valores 0,9266, 0,9250 e 0,9511 para as bases de linfoma, colorretal e mama, respectivamente. O estudo desse método é especialmente relevante no contexto da normalização de coloração, pois ele busca preservar as estruturas histológicas sem introduzir artefatos visuais, além de adaptar-se automaticamente à variabilidade entre as imagens, oferecendo uma alternativa robusta e biologicamente coerente.

O trabalho de (Pérez-Bueno et al. 2022) apresentou uma modelagem Bayesiana e uma estrutura de deconvolução de cores de inferência baseada no algoritmo de decomposição de valor K -singular (do inglês, *K-Singular Value Decomposition* (K-SVD)). O K-SVD é um algoritmo guloso para aprendizado de dicionário e representação esparsa de sinais. O dicionário e a representação esparsa são a matriz de aparência dos corantes e o mapa de densidades dos corantes, respectivamente. Diferentemente de técnicas com uma matriz fixa de cores para os corantes, o *Bayesian K-SVD* (BKSVd) utiliza modelagem probabilística Bayesiana, permitindo que os espectros cromáticos sejam inferidos dinamicamente. Para a avaliação dos dois métodos de normalização propostos, BKSVd e *Empirical Bayesian K-SVD* (EBKSVd), foram utilizadas imagens de duas bases públicas: o *dataset* WSSB, com imagens H&E de tecidos de mama, cólon e pulmão, e o *dataset* CAMELYON17, desenvolvido para detecção de metástase de câncer de mama. Para avaliar os métodos, os autores extraíram as métricas *Peak Signal-to-Noise Ratio* (PSNR)

e *Structural Similarity Index Measure* (SSIM), e compararam com outros trabalhos da literatura, obtendo os valores 19,29 e 19,54 para PSNR, e 0,8594 e 0,8735 para SSIM. Nessa comparação realizada, os métodos BKSVD e EBKSVD não obtiveram as maiores métricas, mas obtiveram os melhores resultados visuais, segundo os autores, produzindo resultados de alta qualidade e minimizando as variações de cor e mantendo diferenças claras entre ambas as colorações.

3.2 Segmentação de Núcleos em Imagens H&E de Cavidade Oral

As técnicas de aprendizado profundo aplicadas à segmentação conseguem generalizar as estruturas nas imagens enquanto preservam seus dados posicionais, o que resulta em representações eficazes e de alta precisão. Por esse motivo, o uso desses modelos na segmentação de imagens médicas tem se destacado cada vez mais nos últimos anos (Nasir et al. 2023). A área de pesquisa sobre segmentação de imagens é bastante abrangente, incluindo diversos algoritmos e cenários de uso. Porém, neste trabalho, serão considerados apenas estudos realizados sobre imagens de tecidos de displasia, a fim de obter uma comparação mais precisa nesse contexto histológico.

O trabalho de (A. B. Silva et al. 2019) propôs um método de segmentação automática de núcleos celulares em tecidos epiteliais orais com displasia utilizando CNNs. O modelo baseia-se na arquitetura Mask R-CNN com *backbone* pelo modelo ResNet-50, conhecida por sua capacidade de realizar segmentação por instância com alta precisão. Foi utilizada transferência de aprendizado a partir da base *ImageNet*, e treinamento da segmentação com imagens de displasia bucal em tecidos de camundongos expostos a um carcinógeno. Para isso, foram utilizadas máscaras manuais de núcleos, processadas com aumento de resolução e anotadas por especialistas. A segmentação foi refinada por um estágio de pós-processamento que emprega operações morfológicas como dilatação, erosão e preenchimento de buracos a fim de reduzir falsos positivos e negativos. A abordagem foi avaliada com base em métricas como acurácia, coeficiente *Dice* e correspondência com o padrão-ouro, alcançando os valores 89,52, 0,76 e 0,84, respectivamente, atingindo um desempenho superior em comparação com os métodos K-means, EM-GMM e SegNet. Os resultados indicaram que o modelo foi eficaz mesmo em cenários desafiadores, como a segmentação de núcleos em regiões com displasias severas, contribuindo significativamente para o avanço de sistemas de diagnóstico assistido por computador em histopatologia.

Um estudo relevante na área de segmentação de imagens histológicas foi apresentado por (D. F. D. Santos et al. 2020a), que investigou os impactos da transformação de espaço de cor na segmentação de núcleos displásicos utilizando CNNs em uma base de tecidos de línguas de camundongos previamente diagnosticados com displasia bucal. O método proposto consiste em aplicar diferentes transformações de espaço de cor (incluindo

RGB, HSV e CIE L*a*b*) em imagens histológicas de tecidos orais, para então treinar redes totalmente convolucionais de forma independente em cada variação. O objetivo foi avaliar como a representação da imagem em diferentes espaços influencia a qualidade da segmentação dos núcleos celulares, que é uma etapa crítica em análises patológicas assistidas por computador. Os resultados demonstraram que o espaço de cor CIE L*a*b* obteve o melhor desempenho em termos de acurácia, com 0,882, sugerindo que componentes relacionados à luminância (estrutura da imagem) são mais relevantes para essa tarefa do que componentes de cromaticidade (cor). Apesar disso, o espaço de cor RGB obteve o maior valor para a métrica de coeficiente *Dice*, com 0,82. Este trabalho destacou a importância da escolha do espaço de cor como uma etapa estratégica em pipelines de segmentação baseados em aprendizado profundo.

O estudo de (D. F. D. Santos et al. 2020b) propôs uma abordagem para a segmentação automatizada de núcleos celulares em imagens histológicas de tecidos orais displásicos, utilizando uma versão modificada da arquitetura U-Net. Uma das principais contribuições do trabalho está na análise do impacto de diferentes técnicas de normalização de cores no desempenho da segmentação. Embora a normalização tenha sido aplicada para reduzir variações de iluminação e coloração nas imagens, os autores concluíram que ela não trouxe benefícios significativos para a precisão da segmentação dos núcleos. Em contraste, as modificações na arquitetura U-Net, como a normalização em lote e filtros convolucionais adaptados, foram cruciais para o sucesso do modelo. A rede apresentou uma segmentação com acurácia de 0,879 e coeficiente *Dice* de 0,820, mesmo sem a necessidade de normalização das cores. Esses resultados sugerem que, para este caso específico de segmentação histológica, a modificação arquitetural foi mais relevante do que o pré-processamento das imagens, destacando a importância da adaptação da rede às características das imagens biomédicas.

No trabalho de (A. B. Silva et al. 2020), foi proposta uma abordagem baseada em CNNs para a segmentação de núcleos celulares em imagens histológicas de displasias epiteliais orais, empregando a arquitetura Mask R-CNN aliada a técnicas de normalização de cor. Os autores realizaram uma análise comparativa entre diferentes otimizadores, como Adam, SGD (do inglês, *Stochastic Gradient Descent*) e Nadam, além de avaliarem o impacto da normalização de cor no desempenho do modelo. Os resultados mostraram que o melhor desempenho foi alcançado com o otimizador Nadam aplicado diretamente às imagens originais, sem qualquer normalização de cor, atingindo uma acurácia média de 88,7%, sensibilidade de 76,2% e especificidade de 94,2%. Embora a normalização de cor tenha apresentado algum impacto positivo em dois dos métodos de segmentação testados, o estudo demonstrou que, de forma geral, essa técnica resultou em queda de desempenho nos demais casos, além de comprometer a integridade das estruturas na região dos núcleos. Dessa forma, os autores destacaram que a aplicação indiscriminada da normalização de cor pode ser prejudicial, reforçando a necessidade de uma escolha criteriosa e dependente

do contexto específico em tarefas de segmentação com CNNs.

No estudo de (A. B. Silva et al. 2022), os autores propuseram uma abordagem baseada em aprendizado profundo para a segmentação de núcleos celulares em imagens oriundas de tecidos com displasia epitelial oral. Utilizando a arquitetura Mask R-CNN, foram segmentados 9.155 núcleos em 296 regiões de imagens, com acurácia entre 88,92% e 90,35% para os diferentes níveis de gravidade da lesão. A metodologia incluiu um estágio de pós-processamento com operações morfológicas para remoção de falsos positivos e negativos, aprimorando a qualidade da segmentação. As regiões segmentadas serviram de base para a extração de 23 atributos morfológicos e texturais, posteriormente usados na classificação do grau de displasia. O estudo destacou a eficácia do uso de redes de segmentação por instância em contextos histopatológicos, evidenciando seu potencial na construção de sistemas computacionais de apoio a diagnósticos.

(A. B. Silva et al. 2024b) apresentaram o *OralEpitheliumDB*, um conjunto de dados público e anotado de imagens histológicas de displasia epitelial oral extraídas de camundongos, desenvolvido para auxiliar o treinamento e avaliação de modelos de segmentação e classificação. Composto por 456 imagens coradas com H&E, o banco de imagens inclui versões originais e normalizadas das imagens, além de segmentações de referência. Para a segmentação de núcleos celulares, os autores empregaram a arquitetura U-Net com *backbone* ResNet-50. Ao avaliarem o impacto da normalização de cor, observaram que, embora o método de (Reinhard et al. 2001) tenha proporcionado melhorias pontuais, os resultados foram inconsistentes e sensíveis à imagem de referência escolhida, atingindo acurácia de 85,44%, que, por sua vez, obteve o melhor resultado sem considerar a normalização, com 90,56%. Assim, concluíram que a normalização pode beneficiar ou prejudicar a segmentação dependendo do contexto, reforçando a necessidade de uma análise criteriosa da combinação entre a técnica de normalização e a arquitetura utilizada.

3.3 Considerações finais

Pelos desafios identificados em cada trabalho citado, verifica-se que a aplicação da normalização de cores, no contexto das imagens histológicas H&E, não tem uma metodologia consolidada, apresentando trabalhos em que sua utilização pode afetar os resultados tanto negativamente quanto positivamente. Desse modo, denota-se que há espaço para estudos comparativos na área de normalização de imagens histológicas H&E. A utilização do método de aprendizado de dicionários SNMF possui uma importante premissa ao considerar conceitos biológicos dos tecidos e dos corantes. Uma investigação dessas estratégias em etapas de processamento de imagens, como na segmentação, uma das etapas primordiais para extração e classificação aplicadas em sistema CAD, ainda é um ponto em aberto na literatura.

Metodologia

Neste capítulo, são apresentadas as abordagens investigadas para normalização dos corantes H&E e o método para segmentação de núcleos em imagens histológicas da cavidade oral. Também são apresentados o banco de imagens empregado na investigação dos modelos propostos, a interpretação das características dentro das camadas convolucionais do modelo de segmentação dos núcleos e as métricas quantitativas para análise dos resultados.

4.1 Introdução

Na Figura 14 é apresentado um diagrama em blocos que representa a metodologia desenvolvida neste trabalho. Essas etapas foram divididas em: i) investigação de métodos de normalização de cores baseados em aprendizado de dicionário com o banco de imagens histológicas de displasia; ii) modelo de segmentação baseado em uma aproximação da U-Net para segmentação de núcleos nas imagens, treinado e testado a partir das imagens normalizadas e não normalizadas; iii) interpretação das características nas camadas convolucionais do modelo CNN a partir da metodologia ProtoSeg.

Os algoritmos propostos foram desenvolvidos usando a linguagem de programação Python com a biblioteca Keras. Os experimentos foram executados em um computador Windows 10 com processador Intel(R) Core(TM) i7-9700K, 16GB de memória RAM e GPU Nvidia GeForce RTX 3060 com 8GB de VRAM.

4.2 Banco de Imagens

O banco de imagens aplicado neste estudo é composto por tecidos obtidos da região da cavidade oral de camundongos. Essas imagens foram coletadas por meio de experimento *in vivo* devidamente aprovado pelo Comitê de Ética da UFU na utilização de animais sob número de protocolo 038/09, realizado no período de 2009 a 2010. Essas imagens foram empregadas nos trabalhos apresentados por (A. B. Silva et al. 2019), (D. F. D.

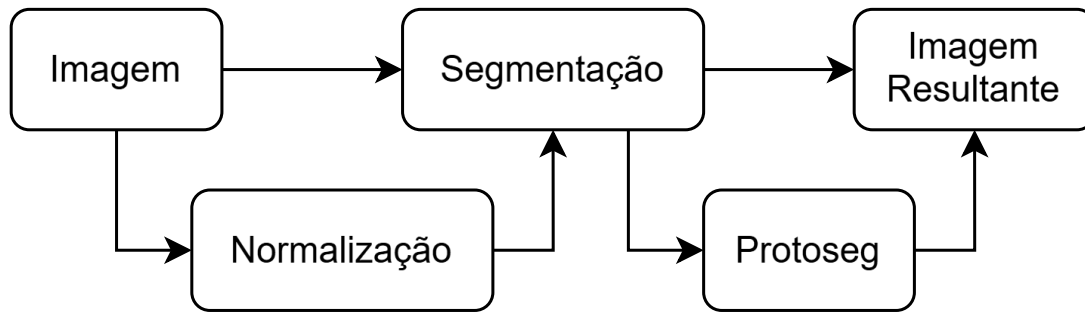


Figura 14 – Diagrama das etapas da metodologia proposta para normalização, segmentação e interpretação das características.

d. Santos et al. 2020a), (A. B. Silva et al. 2022), (A. B. Silva et al. 2023), (A. B. Silva et al. 2024c) e (A. B. Silva et al. 2024b), e disponibilizadas publicamente no GitHub, <https://github.com/LIPAIGroup/OralEpitheliumDB>.

As imagens histológicas foram capturadas por meio do microscópio óptico Leica DM500, com ampliação de $400\times$, e armazenadas em arquivos no formato TIF. As imagens foram adquiridas em modelo de cores RGB, com quantificação de 8 bytes e resolução de 2.048×1.536 pixels. No total, foram digitalizadas 134 lâminas, e as imagens geradas foram segmentadas em regiões de interesse (ROI) com dimensões de 450×250 pixels. Além disso, as imagens foram divididas em classes de acordo com o grau de severidade da lesão: tecido saudável, displasia leve, displasia moderada e displasia severa. Cada classe de tecido possui 114 ROIs, resultando em um total de 456 ROIs. Na Figura 15 são ilustrados exemplos dos tecidos digitalizados e armazenados no banco de imagens em que estão incluídas lesões de displasia leve, moderada e severa, e tecido saudável. A divisão de classes não foi considerada neste estudo, os processos de normalização e segmentação foram executados sobre todas as imagens, sem distinção.

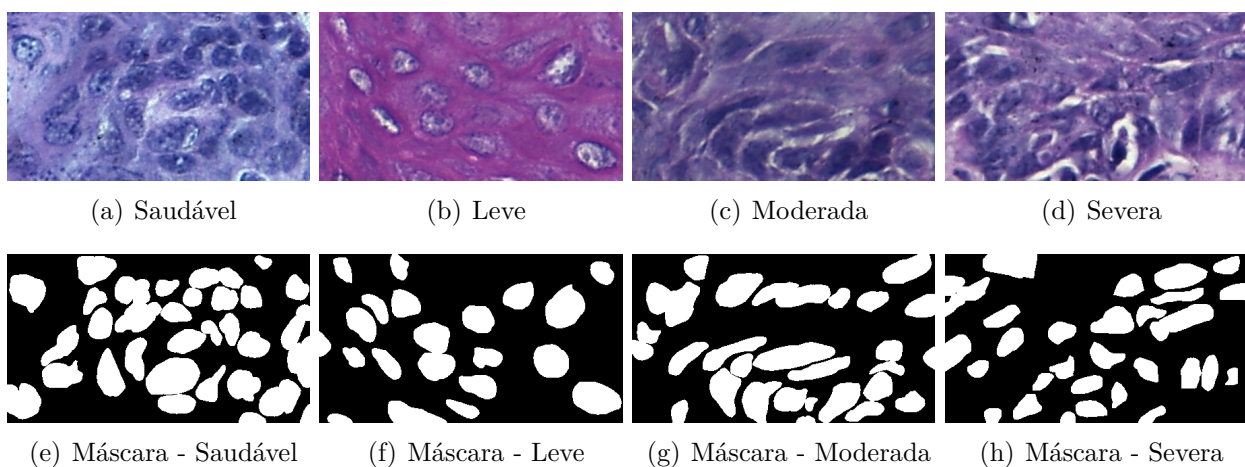


Figura 15 – Tecidos histológicos da cavidade oral com imagem saudável e com lesões displásicas. Na linha superior são mostradas as ROIs e a linha inferior apresenta as máscaras do padrão ouro delineado pelo especialista (Figura adaptada de (A. B. Silva et al. 2024b)).

Como as imagens da base possuíam dimensões 450×250 pixels, como forma de aproveitar a maior quantidade de informações, as seguintes operações foram realizadas: i) as imagens foram redimensionadas para o tamanho 512×284 pixels; ii) foram extraídos os recortes (*crops*) em tamanhos 256×256 pixels para condizer com as dimensões de input da rede CNN. Na Figura 16 é mostrado o procedimento de recorte realizado nas imagens originais. Na Figura 16(b), o quadrante vermelho representa o primeiro *crop*, o quadrante verde representa o segundo *crop*, e o quadrante amarelo representa a informação descartada devido ao ajuste aplicado. As imagens das Figuras 16(c) e (d) mostram os novos formatos das imagens após a operação de recortes. A partir desses *crops*, o conjunto de imagens, formado por 456 ROIs foi aumentado e dividido em conjuntos de treinamento, validação e teste, com 638, 90 e 184 imagens, respectivamente.

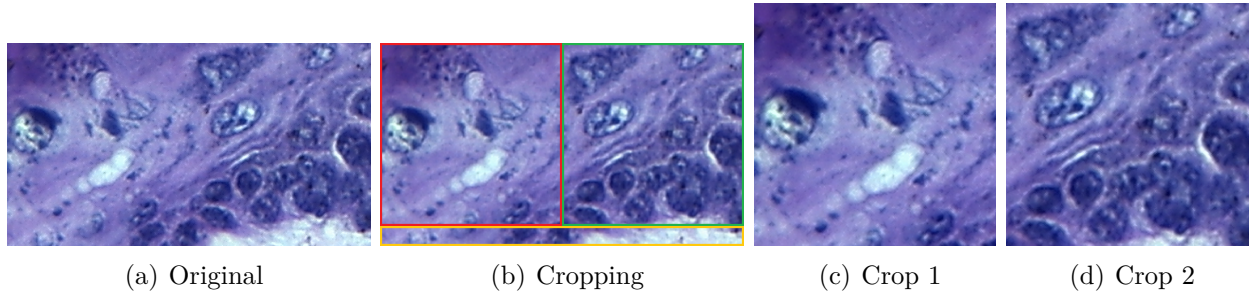


Figura 16 – Exemplo de *crop* realizado na base de imagens para obter imagens de entrada no padrão da arquitetura usada para segmentação.

4.3 Normalização de Corantes de Hematoxilina e Eosina

Nesta etapa foram investigados os métodos de normalização dos corantes H&E baseados em abordagens de aprendizado de dicionário. Para isso, foram consideradas as técnicas: distância Euclidiana (do inglês, *Euclidean Distance* (ED)), divergência de *Renyi*, e divergência de *Kullback-Leibler* (KL) (Devarajan et al. 2016). Esses métodos exploram a estimativa de coeficientes da matriz de aparência de cores dos corantes para alterações de cores nas imagens.

Para implementação das abordagens, as imagens de entrada (referência e original) foram usadas para a estimativa de coeficientes de corante H&E. A lei de *Beer-Lambert* (Equação 2) estabelece que a atenuação do espectro de luz (canais do modelo de cores RGB) incidente sobre amostras coradas é dada pela concentração dos corantes absorvidos pelas estruturas do tecido investigado:

$$I = I_0 \cdot \exp(-W \cdot H). \quad (2)$$

Para uma imagem de entrada I , representada nos canais do modelo RGB distribuídos em $M \times N$, em que M corresponde aos canais e N ao número de pixels da imagem, e I_0 corresponde à intensidade da luz incidente sobre a amostra histológica, são definidas a matriz W com dimensão $M \times R$, onde R pode representar os corantes H&E aplicados sobre a amostra dado por $R = 2$ (hematoxilina e eosina), e o mapa H de dimensão $R \times N$ (Vahadane et al. 2016).

As densidades ópticas relativas dos canais R (I_R), G (I_G) e B (I_B) de uma imagem, respectivamente denotadas por V_R , V_G e V_B , são definidas por:

$$V_R = \log \frac{I_0}{I_R}, \quad (3)$$

$$V_G = \log \frac{I_0}{I_G}, \quad (4)$$

$$V_B = \log \frac{I_0}{I_B}. \quad (5)$$

Com isso, a relação entre as densidades ópticas V , definidas pela junção entre V_R , V_G e V_B , e a lei de Beer-Lambert é estabelecida por:

$$V = W \cdot H. \quad (6)$$

Dada a combinação linear entre os termos W e H , então, uma imagem histológica corada com H&E pode ser representada. A matriz W é composta por coeficientes que especificam como os corantes são representados pelos canais do modelo RGB. Então, o (i, j) -ésimo coeficiente representa o espectro de absorção do j -ésimo corante no i -ésimo canal do modelo RGB. O H descreve a proporção de cada corante representado pelos pixels da imagem (TOSTA et al. 2019a). Dessa forma, o (i, j) -ésimo coeficiente corresponde à quantificação do i -ésimo corante no j -ésimo pixel da imagem.

Os métodos ED, KL e *Renyi* foram explorados neste estudo para estimativa das matrizes W das imagens originais e de referência por meio da ferramenta High-Performance Computing Non-negative Matrix Factorization (hpcNMF). Essa ferramenta foi escolhida por ser uma implementação recente, bem estruturada e de código aberto, o que facilita a reprodutibilidade dos experimentos e evita vieses relacionados a implementações proprietárias. Essa ferramenta também oferece uma ampla variedade de abordagens para aprendizagem de dicionários, permitindo a comparação entre diferentes funções de custo aplicadas à tarefa de normalização de colorações em imagens histológicas coradas com H&E (Devarajan et al. 2015).

Diversas funções podem ser aplicadas para estimar a matriz W de aparência de cores dos corantes H&E. No geral, essas funções de custo são derivadas da divergência KL, como é o caso da Distância Euclidiana (ED), que calcula a diferença quadrática entre os valores, definida por:

$$ED(V\|WH) = \frac{1}{2\sigma^2} \sum_{i,j} (V_{ij} - (WH)_{ij})^2, \quad (7)$$

onde a variável σ^2 representa a variância do ruído, assumindo que os dados estão sujeitos a perturbações aleatórias com distribuição Gaussiana. Os índices i e j indicam a posição dos elementos dentro das matrizes, correspondendo às linhas e colunas, respectivamente. O somatório percorre todos esses elementos para calcular o erro total entre os valores observados e os estimados.

Já a divergência de *Renyi* está relacionada à probabilidade de Poisson do produto das matrizes W e H gerar a matriz V , a qual é definida por:

$$D_\alpha(V\|WH) = \frac{1}{\alpha - 1} \sum_{i,j} \left[V_{ij}^\alpha (WH)_{ij}^{1-\alpha} - \alpha V_{ij} - (1 - \alpha)(WH)_{ij} \right], \quad (8)$$

em que um parâmetro ajustável α ($\alpha \neq 1$) permite controlar a sensibilidade do modelo a discrepâncias maiores ou menores.

A divergência de Kullback-Leibler KL é um caso especial da divergência de *Renyi* quando α tende a 1 ($\alpha \rightarrow 1$), resultando em:

$$KL(V\|WH) = \sum_{i,j} \left[V_{ij} \log \left(\frac{V_{ij}}{(WH)_{ij}} \right) - V_{ij} + (WH)_{ij} \right]. \quad (9)$$

A ferramenta hpcNMF permite o controle de parâmetros específicos para a execução dos algoritmos NMF. Além da seleção dos algoritmos de fatoração (ED, KL e *Renyi*), outros parâmetros podem ser ajustados para otimizar o desempenho do método selecionado. O parâmetro de número de fatores k determina quantos fatores serão extraídos. Para determinar a matriz de aparência de cores W dos corantes H&E, o parâmetro k deve ser atribuído ao valor 2. A ferramenta oferece a possibilidade de especificar o número de iterações n em que o algoritmo NMF realizará para convergir para uma solução. Além disso, o parâmetro de número de execuções c determina quantas vezes o processo de fatoração será repetido com diferentes inicializações aleatórias da matriz W . Os valores dos parâmetros escolhidos nesse trabalho foram $c = 1$ e $n = 100$, os quais foram baseados nos experimentos realizados no estudo apresentado por (A. F. Q. d. Silva 2024).

De forma semelhante ao estudo de (Tosta et al. 2024), o método também aplica a estimativa do mapa H , o qual é tratado como um problema de mínimos quadrados linear $l1$ -regularizado. Então, nesta fase, o algoritmo LARS baseado em *Cholesky* foi utilizado (Julien Mairal et al. 2010) com implementação pela ferramenta SPAMS (J. Mairal et al. 2017).

Tendo estabelecido as estimativas de W e H , o processo de normalização ocorre na imagem de entrada. Então, o mapa calculado de H da imagem original ($H_{original}$) é combinado com a matriz W da imagem de referência ($W_{referência}$). Neste processo, uma imagem de referência deve ser escolhida por especialista como uma imagem que apresente

boas condições de coloração dos corantes nos tecidos. Para que os valores dos mapas H das imagens, original e referência, estejam contidos em um mesmo intervalo, a Equação 10 foi empregada para o ajuste desses valores, conforme descrito em (Vahadane et al. 2016):

$$H_{original}^{norm}(j, :) = \frac{H_{original}(j, :)}{H_{original}^{RM}(j, :)} H_{referência}^{RM}(j, :), j = 1, \dots, R, \quad (10)$$

onde $H_i^{RM} = RM(H_i) \in R^{R \times 1}$, $i = \{original, referência\}$ e RM corresponde ao pseudo máximo de cada linha do mapa H em 99%.

4.4 Estratégia de Aumento de Dados

Em imagens médicas, o número de amostras é normalmente pequeno para investigação sobre modelos de aprendizagem profunda. Esse cenário pode, em muitos casos, não ser suficiente para treinar adequadamente o modelo, podendo gerar *overfitting* durante o treinamento do modelo, o qual é caracterizado pelo distanciamento das curvas de treinamento (azul) e validação (verde) (ver Figura 17). Há casos em que abordagens de transferência de aprendizado podem não ser adequadas para domínios que envolvam dados histológicos, pois parte dos pesos aprendidos por essa estratégia é obtida por uma base generalizada e não desenvolvida para a etapa de segmentação.

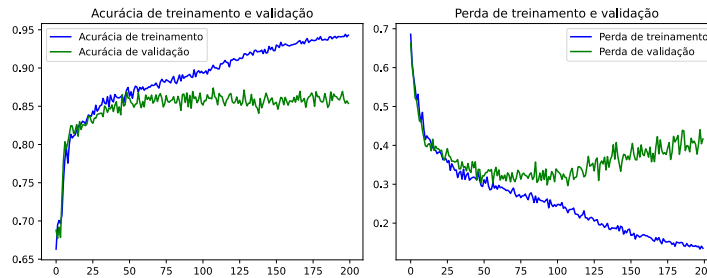


Figura 17 – Gráfico de treinamento e validação de uma rede CNN com *overfitting*.

De modo a solucionar o problema de *overfitting*, neste estudo foi aplicada a técnica de aumento de dados no banco de imagens. A técnica escolhida é chamada RCAug (*Random Composition Augmentation*), apresentada por (D. F. D. Santos 2022), a qual realiza uma composição aleatória de transformações sobre as imagens de entrada e seus respectivos rótulos, promovendo uma maior variabilidade no conjunto de treinamento. A cada execução do método, cada transformação tem uma probabilidade empiricamente definida de 50% de ser aplicada, permitindo que diferentes combinações sejam geradas de forma dinâmica. Essa abordagem garante que a rede neural seja exposta a uma ampla gama de variações sem aumentar significativamente os requisitos de armazenamento, contribuindo para a robustez do modelo e melhorando sua capacidade de generalização em diferentes cenários.

As transformações geométricas foram empregadas para minimizar erros de posicionamento e orientação nas imagens histológicas, tornando o modelo mais resistente a variações espaciais nos dados. Entre essas transformações, destacam-se rotações, transposições e espelhamentos (*flips*) nos eixos horizontal e vertical, garantindo que o modelo não dependa de uma configuração específica das estruturas na imagem. Além disso, para lidar com artefatos indesejados, como dobras no tecido, distorções ópticas, borrões e inconsistências de coloração causadas por diferentes técnicas de coloração e variações nas condições de iluminação, foram aplicadas transformações específicas para reduzir esses efeitos e aprimorar a qualidade visual das imagens processadas.

A operação do método RCAug foi aplicada 3 (três) vezes sobre a base de imagens originais de treinamento, que continha 638 *crops*. Dessa forma, ao aproveitar também as imagens originais, foi possível quadruplicar a quantidade total de dados de treino e as respectivas máscaras para a operação de treinamento do modelo, resultando em 2552 imagens para treinamento.

4.5 Segmentação de Núcleos Celulares nas Imagens Histológicas

A aproximação desenvolvida foi baseada no modelo de segmentação U-Net para a segmentação dos núcleos celulares das imagens histológicas. Esse modelo possui estruturas denominadas “*encoder*”, “*bottleneck*” e “*decoder*”, que representam grupos de camadas da rede. Na entrada do modelo foram empregadas as imagens com canais de cores RGB em dimensões de 256×256 pixels. A arquitetura adaptada é apresentada na Figura 18. Baseada na arquitetura U-Net, camadas de *dropout* espacial (setas marcadas em amarelo) foram adicionadas para remover o comportamento de *overfitting*.

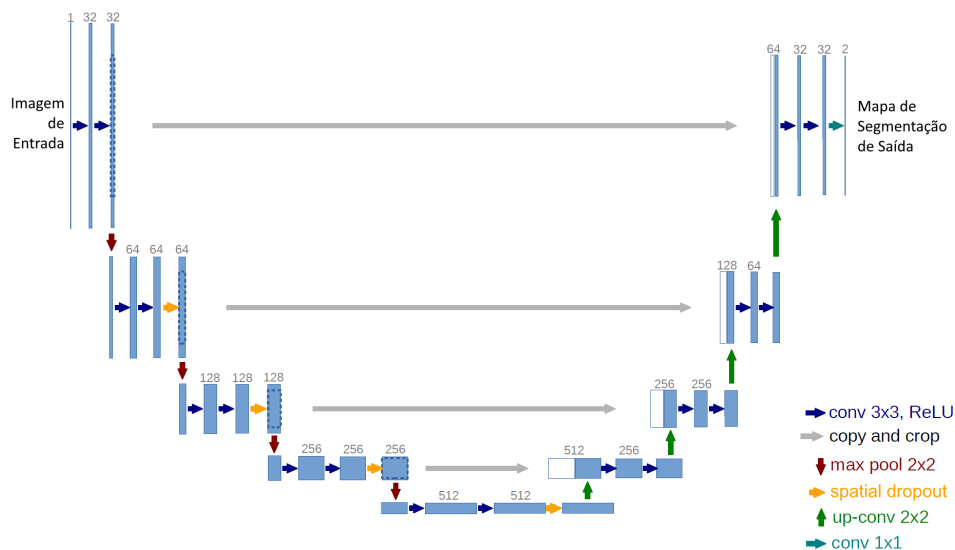


Figura 18 – Arquitetura da rede CNN apresentada neste trabalho.

Apesar da etapa de aumento de dados ser prevista e realizada no estudo, foi observado empiricamente a ocorrência de *overfitting*. A fim de remover esse comportamento no modelo, alguns métodos e modificações na rede original foram testados. Métodos de regularização como *dropout*, *dropout* de Monte Carlo, *dropout* espacial e L2, assim como a associação entre eles nos blocos, foram analisados, com base nos estudos de (Goodfellow et al. 2016; Tompson et al. 2015; Srivastava et al. 2014; Gal et al. 2016). Na Figura 19 é mostrado o comportamento do modelo avaliado sobre as diferentes técnicas e parametrizações de regularização no modelo.

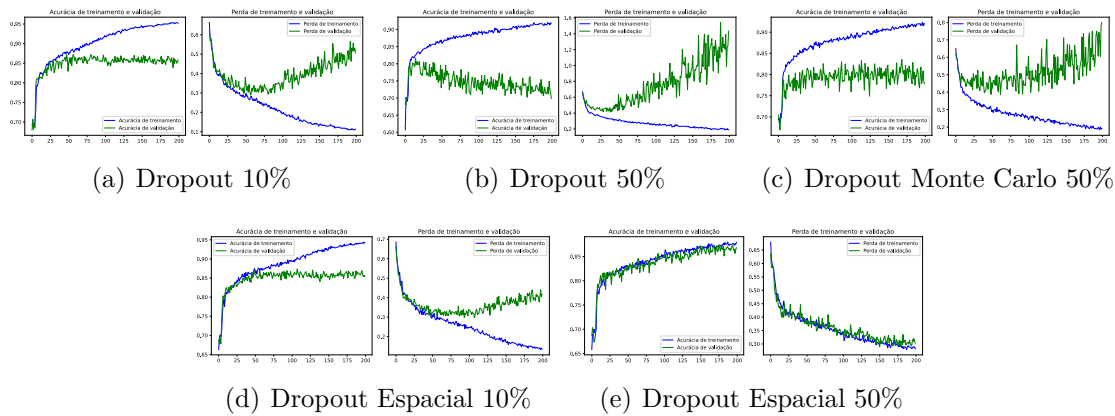


Figura 19 – Gráficos de treinamento utilizando os diferentes abordagens de regularização analisadas no modelo.

A regularização *dropout* clássica é uma técnica usada em CNNs para evitar que o modelo dependa excessivamente de certos padrões nos dados de treinamento. Isso ajuda a melhorar a capacidade do modelo de fazer previsões para novos dados, aumentando a aleatoriedade do treinamento e reduzindo o *overfitting*. Esse método funciona desativando aleatoriamente canais inteiros de uma imagem ou de um mapa de características durante o treinamento. Isso significa que, em cada passagem, alguns canais são completamente ignorados. Essa abordagem busca auxiliar a rede a aprender a reconhecer padrões sem depender excessivamente de um único canal específico. Já o *dropout* de Monte Carlo estende essa ideia para a fase de inferência, permitindo obter distribuições de saída e medir incertezas, sendo especialmente útil em aplicações críticas. Por outro lado, a regularização L2 penaliza diretamente pesos grandes na função de perda, incentivando soluções mais simples e estáveis. Enquanto o *dropout* é mais indicado para redes densas, o L2 é amplamente utilizado em CNNs em todas as camadas, e o *dropout* Monte Carlo é aplicado quando é necessário quantificar a incerteza nas previsões. Todas essas técnicas ajudam a tornar o modelo mais robusto e generalizável, mas a escolha entre elas depende da arquitetura da rede e do tipo de dados com os quais se está trabalhando (Srivastava et al. 2014; Goodfellow et al. 2016; Gal et al. 2016).

No estudo dos processos aplicados, a utilização de *dropout* espacial proporcionou os modelos mais adequados para avaliação. O *dropout* espacial é uma variação do *dropout* que

também desativa canais inteiros, mas de uma forma mais consistente. Em vez de escolher aleatoriamente canais diferentes para cada posição espacial, ele garante que um mesmo canal seja desligado por completo para toda a amostra. Isso evita que a rede se baseie excessivamente em determinados mapas de características, incentivando-a a considerar mais informações globais na imagem ao ajudar a promover a independência entre mapas de características (Tompson et al. 2015).

Uma investigação foi realizada sobre as camadas, e o *dropout* espacial de 50%, que obteve melhores resultados empíricos, foi adicionado após a quarta, sexta e oitava camadas convolucionais, no *encoder*, e após a décima camada convolucional, no *bottleneck*. Não foram adicionadas camadas no grupo das camadas de *decoder* pois, nesta etapa, a rede está “reconstruindo” as informações de modo a gerar a imagem resultante, não sendo necessário eliminar informações que podem ser valiosas para essa reconstrução. Esse modelo foi treinado com o otimizador Adam, com taxa de aprendizado (ou *learning rate*) de 0,0001, função de perda de Entropia Cruzada Binária (ou *Binary Cross-Entropy*), tamanho do lote (ou *batch size*) de 16 imagens e 200 épocas de treinamento, com *input* de imagens de tamanho 256×256 pixels. Os modelos foram treinados na abordagem “*from scratch*”, sem nenhum modelo pré-treinado como base.

4.6 Interpretação das Características no Modelo de Segmentação

Uma estratégia para entender o comportamento e melhorar o desempenho do modelo é analisar os dados gerados durante seu treinamento. Essa estratégia também pode ser usada para entender o comportamento da normalização sobre o processo de segmentação. Para tanto, os autores em (S. He et al. 2023) apresentaram um método de interpretação das características, denominado *Prototype Segmentation* (ProtoSeg), o qual permite analisar as características (*deep features*) em imagens com abordagem semântica.

Esse método explora a interpretação de forma a avaliar como a atribuição de uma métrica à função de perda de uma CNN auxilia no avanço do modelo durante o treinamento. Além disso, permite a extração de métrica em cada camada convolucional da arquitetura de forma a entender como a segmentação atua na imagem entre a entrada e saída, e ainda apresenta o conceito de mapa de capacidade de segmentação (do inglês, *Segmentation Ability Map* (SAM)), em que um mapa de segmentação é calculado a partir dos protótipos positivos e negativos gerados a partir dos mapas de características de uma camada e da máscara de segmentação predita.

O procedimento para obtenção do SAM ocorre da seguinte forma: dado $\mathbf{X} \in \mathbb{R}^{H \times W \times C}$ uma matriz de características extraídas do modelo, onde H e W são as dimensões espaciais da imagem e C é o número de canais e $\mathbf{P} \in [0, 1]^{H \times W}$ a máscara de segmentação predita, indicando a probabilidade de um pixel pertencer à classe segmentada. Os protótipos são

obtidos por vetores de tamanho C calculados como a média ponderada das características dentro e fora da região segmentada. O *protótipo positivo* representa a média das características nos pixels que pertencem à região segmentada, definido como:

$$\mathbf{p}_{\text{pos}} = \frac{\sum_{i,j} \mathbf{X}_{i,j} \cdot P_{i,j}}{\sum_{i,j} P_{i,j}}. \quad (11)$$

Neste estudo, essas regiões positivas foram determinadas como as regiões dos núcleos celulares. O *protótipo negativo* representa a média das características nos pixels fora da região segmentada (nesse contexto, o fundo da imagem), definido de maneira similar, utilizando a máscara complementar $\mathbf{R} = 1 - \mathbf{P}$:

$$\mathbf{p}_{\text{neg}} = \frac{\sum_{i,j} \mathbf{X}_{i,j} \cdot R_{i,j}}{\sum_{i,j} R_{i,j}}. \quad (12)$$

Para cada pixel da imagem, aplicou-se a distância quadrática negativa em relação aos protótipos:

$$\mathbf{d}_{\text{pos}} = -\sum (\mathbf{X} - \mathbf{p}_{\text{pos}})^2, \quad (13)$$

$$\mathbf{d}_{\text{neg}} = -\sum (\mathbf{X} - \mathbf{p}_{\text{neg}})^2. \quad (14)$$

As distâncias calculadas foram concatenadas em um vetor $\mathbf{d} = [\mathbf{d}_{\text{neg}}, \mathbf{d}_{\text{pos}}]$, os quais passam por uma ativação *softmax* para obter as probabilidades finais de segmentação:

$$\mathbf{P}_{\text{final}} = \text{softmax}(\mathbf{d}). \quad (15)$$

A partir da matriz de probabilidades geradas pelo método *softmax*, a métrica SAM pode calcular a matriz binária representando as probabilidades de cada pixel (1 se a chance positiva é maior que a negativa e 0 caso contrário), gerando uma máscara de segmentação. Dessa forma, o método Protoseg permite refinar o modelo com um mecanismo baseado em protótipos construídos durante o processo de treinamento. Ao considerar a métrica da média das características nas regiões dentro e fora do objeto segmentado, o método pode melhorar a segmentação, tornando-a mais consistente com as características extraídas da imagem. Desta forma, neste estudo, foi possível avaliar o comportamento e a influência da normalização nas camadas, assim como explorar um aprimoramento do modelo com essa estratégia.

Com a informação SAM, é possível utilizar a métrica *Dice* para calcular a habilidade de segmentação desse mapa intermediário, gerando o valor denominado taxa de habilidade de segmentação (do inglês, *SA Score*). Outras métricas podem ser usadas para calcular esse *SA Score*, mas a medida *Dice* foi escolhida por ser bem estabilizada em diversos estudos da literatura para a etapa de segmentação (TOSTA et al. 2019b; D. F. D. d. Santos et al. 2020a; A. B. Silva et al. 2024b).

A Figura 20 mostra a estrutura geral da metodologia ProtoSeg para geração do *SA Score*. A partir do mapa/máscara intermediária gerado por uma camada da rede e pela saída final da rede, pode-se calcular o *Segmentation map* utilizando o método ProtoSeg. A função *Dice* é aplicada entre o SAM e a máscara do padrão-ouro, assim é calculado o *SA Score* daquela camada.

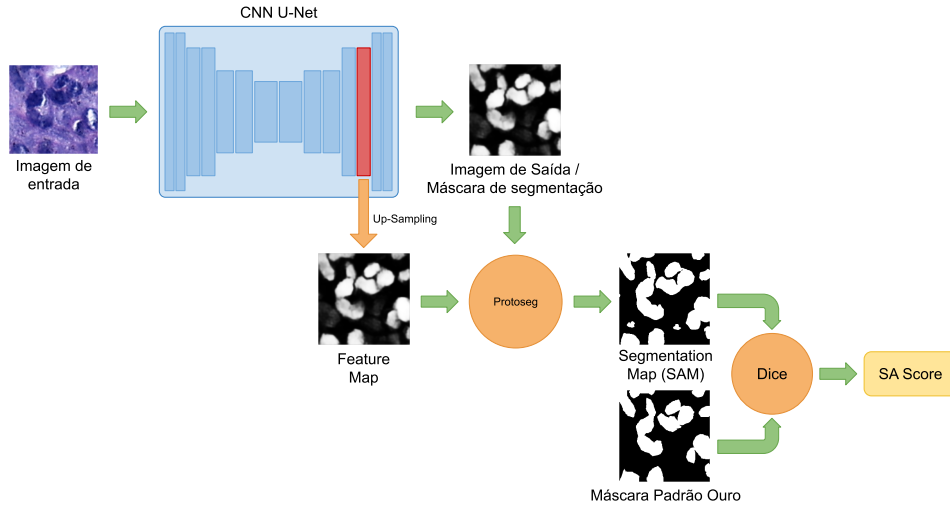


Figura 20 – Estrutura geral da metodologia ProtoSeg para geração do *SA Score*.

O método ProtoSeg pode ser empregado em dois contextos diferentes, denominados Protoseg Offline, quando aplicado nas características de uma CNN treinada, e Protoseg Online, quando o cálculo do *SA Score* é considerado na função de perda da rede durante o processo de treinamento, podendo melhorar o modelo. Neste estudo, as duas abordagens foram consideradas nos experimentos realizados.

4.7 Métricas e critérios de avaliação

Ao realizar os processos e operações relacionadas a um estudo, é necessário utilizar métricas para validar se as operações realizadas são coerentes ao esperado. Para avaliar os processos de normalização de cores e garantir que as transformações preservem informações visuais relevantes, foram consideradas as métricas PSNR, SSIM, *Feature Similarity Indexing Method* (FSIM) e QSSIM.

A métrica PSNR é baseada em erros de pixel e mede a relação entre o sinal original e o ruído introduzido na imagem transformada. Essa métrica é utilizada devido à sua simplicidade e eficiência computacional. Valores mais altos de PSNR indicam menor distorção, sugerindo que a imagem transformada está mais próxima da original, sem adição de ruídos (Setiadi 2021).

A medida SSIM considera estruturas locais, contraste e luminância para mensurar a similaridade entre imagens. Diferente da PSNR, que avalia apenas erros absolutos, a SSIM foca na preservação das relações estruturais dentro da imagem. Isso é feito ao

decompor a imagem em três componentes principais: luminância, contraste e estrutura, comparando-os separadamente para obter uma medida composta de similaridade. O valor da métrica varia entre 0 e 1, onde 1 indica imagens idênticas. A aplicabilidade da SSIM na normalização de cores é significativa, pois permite analisar se a transformação de cores mantém a estrutura e os detalhes da imagem original. Por levar em conta a maneira como os seres humanos percebem diferenças visuais, a SSIM é eficaz na avaliação da qualidade de imagens alteradas, evitando que pequenas mudanças insignificantes em valores de pixel distorçam a percepção real da fidelidade da imagem (Gourdeau et al. 2022).

A métrica FSIM é um aprimoramento da SSIM ao incluir a medida de congruência de fase e a magnitude do gradiente, características perceptualmente significativas. A congruência de fase é baseada na ideia de que as informações estruturais mais importantes da imagem estão localizadas em pontos onde múltiplas frequências de onda estão alinhadas. A magnitude do gradiente complementa essa análise ao medir a variação de intensidade nos pixels, capturando bordas e detalhes essenciais. Essa métrica é particularmente vantajosa para análise de imagens normalizadas, pois foca na preservação de detalhes estruturais e bordas, aspectos fundamentais para a percepção visual da qualidade da imagem. Como a normalização de cores pode alterar sutis variações locais, a FSIM oferece uma avaliação mais robusta ao considerar essas transformações e seu impacto na percepção visual geral da imagem, tornando-se uma ferramenta eficaz para validar técnicas de normalização (Zhang et al. 2011).

Finalmente, a QSSIM expande a SSIM para imagens coloridas utilizando a representação *quaternion*, que modela os canais de cores como um todo, ao invés de tratá-los separadamente. Essa abordagem permite que a análise da similaridade entre imagens seja mais precisa, considerando interações entre os canais de cor (normalmente RGB) de forma unificada. Isso resolve uma limitação da SSIM convencional, que trata cada canal de cor individualmente e pode não capturar corretamente a relação entre eles. Essa propriedade torna a QSSIM especialmente útil para processos de normalização de cores, onde as relações entre os canais são frequentemente modificadas para alcançar um melhor equilíbrio tonal ou correção cromática. Como resultado, essa métrica proporciona uma análise mais fiel à percepção humana, sendo sensível a alterações de matiz, saturação e brilho. Seu uso auxilia na escolha de técnicas que minimizem distorções perceptuais e mantenham a fidelidade visual ao longo das transformações aplicadas à imagem (Kolaman et al. 2012).

Para avaliar as operações de segmentação semântica realizadas neste trabalho, foram consideradas as métricas acurácia e *Dice*. Na segmentação de imagens, a acurácia mede a proporção de pixels corretamente classificados em relação ao número total de pixels da imagem, tanto positivos quanto negativos. É uma métrica simples e intuitiva, mas pode ser enganosa em cenários onde há um forte desbalanceamento entre classes, como quando a maioria dos pixels pertence ao fundo. Nesses casos, uma rede pode alcançar alta acurácia apenas classificando o fundo corretamente, sem segmentar bem os objetos

de interesse (Goodfellow et al. 2016; Molnar 2022). A acurácia é dada por:

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN}, \quad (16)$$

com as variáveis sendo Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN).

O Coeficiente *Dice*, também conhecido como *Dice Score* ou Índice de *SørensenDice*, é uma métrica que avalia a similaridade entre a segmentação predita e a segmentação real. Ele é especialmente útil em tarefas onde a área de interesse é pequena em comparação ao fundo, pois é mais sensível a pequenas discrepâncias nas regiões segmentadas. O *Dice* mede o grau de sobreposição entre as predições e os rótulos verdadeiros, sendo calculado pela Equação 17:

$$\text{Dice} = \frac{2 \times VP}{2 \times VP + FP + FN}. \quad (17)$$

Assim, valores próximos de 1 indicam alta similaridade entre a segmentação prevista e a real, enquanto valores próximos de 0 indicam pouca ou nenhuma sobreposição (Ait Skourt et al. 2018).

Experimentos e Análise dos Resultados

Neste capítulo, são apresentados os experimentos e resultados obtidos pelas estratégias investigadas para normalização dos corantes H&E com aprendizado de dicionário. Também são analisados os resultados do modelo de segmentação e interpretação baseado em características aprendidas, obtidos pelas imagens histológicas da cavidade oral.

5.1 Métodos de normalização de corantes baseados em aprendizado de dicionário

Para esta etapa do estudo, foram utilizadas três imagens de referência, selecionadas para investigar os métodos de normalização, em diferentes estratégias de aprendizado de dicionário. O processo de escolha das imagens de referência é de extrema importância para o processo de normalização de imagens, podendo ser realizado somente por um patologista especializado. Desse modo, escolha das imagens foi feita com base em critérios estabelecidos por um especialista em processamento de imagens histológicas, considerando fatores como contraste entre núcleos e fundo da imagem. Além disso, foram consideradas a distinção de tonalidades entre os corantes H&E e a definição das bordas nucleares. Essa abordagem com diferentes imagens de referência foi também investigada em (A. B. Silva et al. 2024a).

As imagens empregadas como referência para investigação dos métodos de normalização H&E são apresentadas na Figura 21. Essas imagens identificadas, como Referência 1 (Ref 1), Referência 2 (Ref 2) e Referência 3 (Ref 3), foram obtidas de diferentes bases de dados. A imagem Ref 1 (Figura 21(a)) foi obtida de um conjunto privado de imagens de tecidos de lesões orais, a qual apresenta características ideais de distribuição dos corantes e definição de bordas nucleares para identificação visual das estruturas. A Ref 2 (Figura 21(b)) é uma imagem de tecido de câncer de mama do *dataset Nuclei Segmentation* (Janowczyk et al. 2016), em que se destacam as tonalidades distintas dos corantes H&E. A imagem Ref 3 (Figura 21(c)) foi obtida no *dataset Histopathology Classifica-*

tion (Kang et al. 2021) e corresponde a um tecido de linfonodo com alto contraste entre estruturas nucleares.

O uso de múltiplas imagens de referência foi adotado com o objetivo de examinar o impacto desse parâmetro na normalização e na segmentação. A escolha do número de imagens de referência foi baseada no estudo de (A. B. Silva et al. 2024a) com o objetivo de investigar diferentes propriedades: a) a Ref 1 apresenta cores menos saturadas; b) a Ref 2 destaca tons bem distintos para ambos os corantes; c) a Ref 3 exibe uma predominância de eosina, diferenciando-se das demais imagens avaliadas.

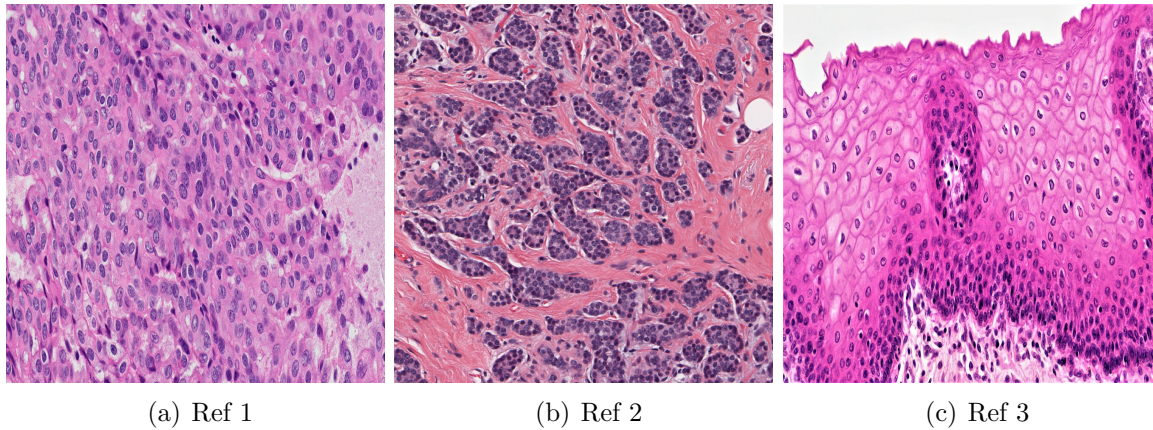


Figura 21 – Imagens de referência utilizadas no processo de normalização dos corantes H&E.

A etapa de normalização foi conduzida com métodos computacionais que usam uma imagem de referência para alterações das cores H&E. Neste estudo foram utilizados os métodos de normalização H&E identificados pelas seguintes nomenclaturas: distância Euclidiana (do inglês, *Euclidean distance*, ED), divergência de Renyi, divergência de *Kullback-Leibler* (KL) (Devarajan et al. 2016), e descritos no capítulo de Metodologia. Também foram aplicados os métodos presentes na literatura descritos no capítulo de Estado da Arte. Esses trabalhos foram identificados por: Li (Xingyu Li et al. 2015), Perez-Bueno (Pérez-Bueno et al. 2022), Tosta (Tosta et al. 2024) e Vahadane (Vahadane et al. 2016). Nas Figuras 22, 23 e 24 são mostradas as imagens obtidas por cada método de normalização das cores dos corantes H&E explorando cada imagem de referência.

Ao comparar visualmente as imagens normalizadas na Figura 22 é perceptível que cada método atingiu um resultado distinto, mostrando que para uma imagem de referência (Ref 1) ocorre uma variação dada pelo método de normalização investigado. Os resultados que mais se assemelharam à Ref 1, a nível de cores, foram os métodos Tosta e Vahadane. Esses métodos mantiveram as regiões claras e áreas escuras da imagem original. Também possibilitaram tratar o contraste, melhorando-o entre os núcleos celulares e as regiões de eosina. Essa informação é destacada na Figura 22(a) pela seta na cor vermelha (eosina) e pela seta na cor verde (núcleo) na ROI da imagem original. Os resultados mostram que o método da divergência de *Renyi* (Figura 22(f)) obteve uma normalização não satisfatória

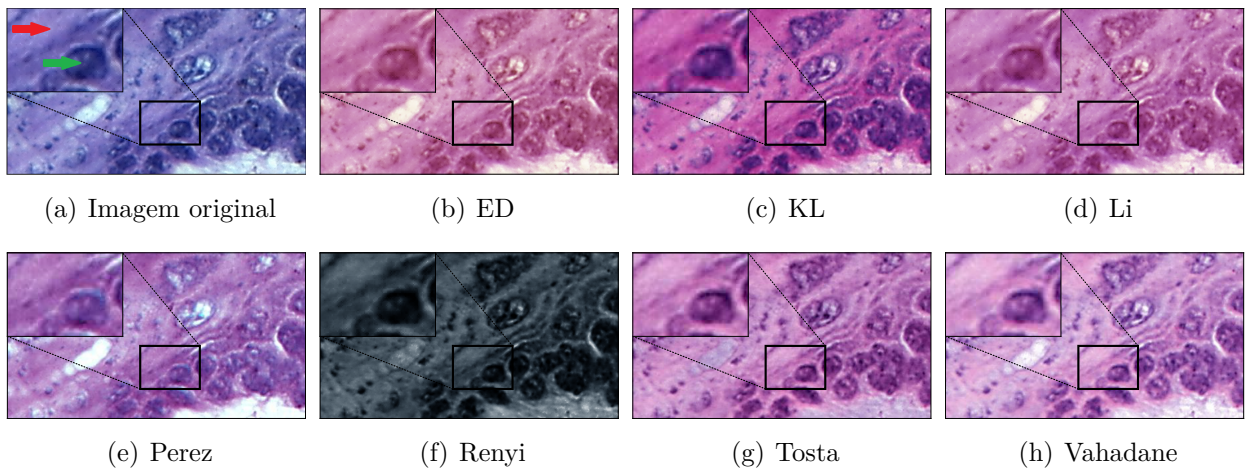


Figura 22 – Resultados da normalização de H&E usando a imagem Ref 1 para investigação dos métodos.

em relação à uniformização dos corantes H&E deixando a imagem com cores próximas ao cinza.

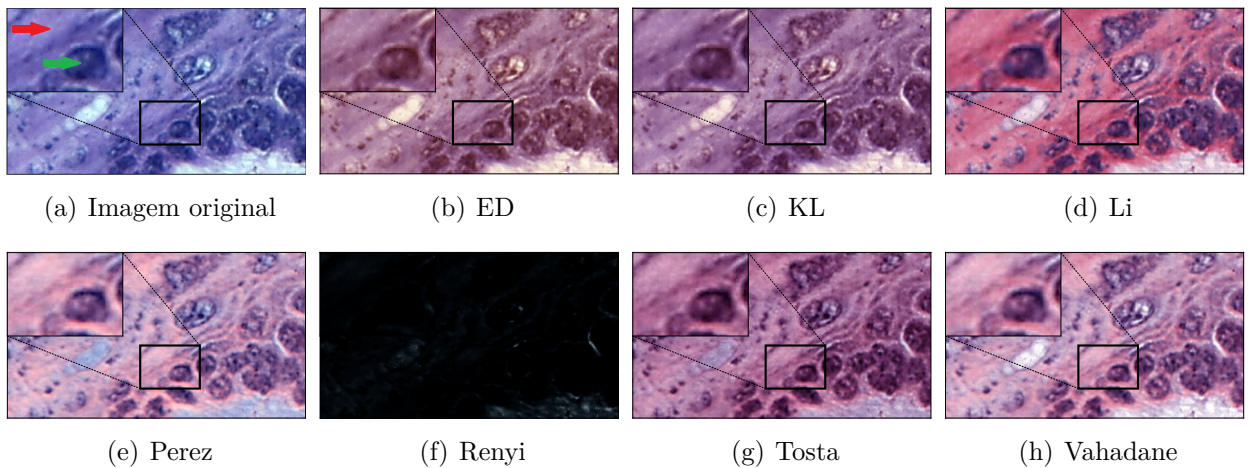


Figura 23 – Exemplos de normalizações de cores realizadas com cada método utilizando a imagem Ref 2.

Para as imagens resultantes da Ref 2 (Figura 23), o método que se destacou na análise visual foi o algoritmo de normalização Li, produzindo uma imagem com cores mais próximas da imagem de referência, mantendo as estruturas claras e escuras dos objetos na imagem, assim como o contraste entre os núcleos e porções coradas predominantemente pela eosina. Os resultados também mostraram que o método da divergência de *Renyi* (Figura 23(f)) obteve uma normalização não satisfatória em relação à uniformização dos corantes H&E, pois parte das intensidades de brilho foi degradada em relação à imagem original.

Para os resultados da normalização utilizando a Ref 3 (Figura 24), observa-se que os melhores resultados visuais ocorreram com os métodos Li, Perez e Vahadane. Esses algoritmos mantiveram as estruturas originais e aumentaram o contraste possibilitando

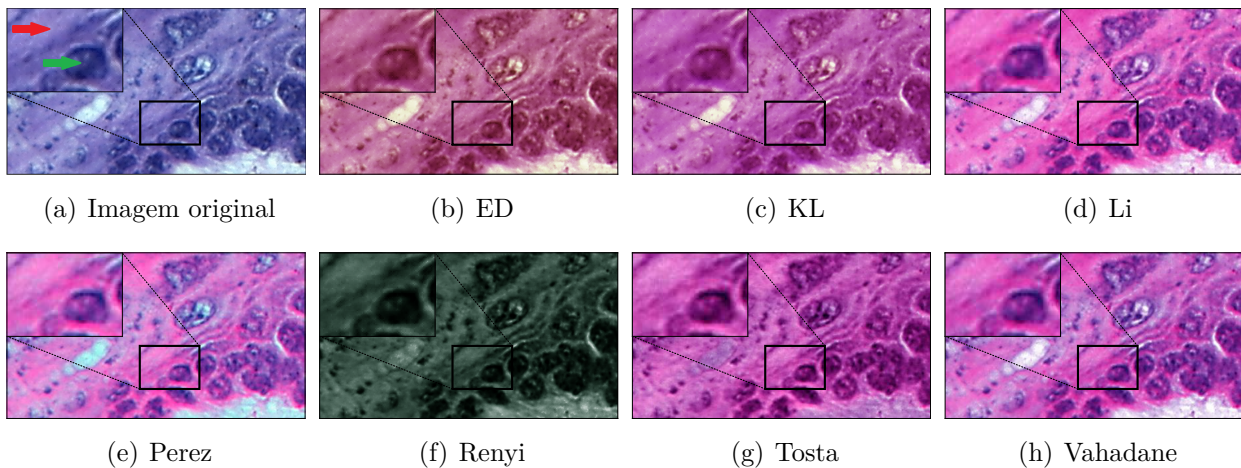


Figura 24 – Investigação da normalização dos corantes realizada com cada método utilizando a imagem Ref 3.

Tabela 1 – Métricas de normalização dos métodos estudados com uso da imagem Ref 1.

Método	FSIM	PSNR	QSSIM	SSIM
ED	0,897	43,244	0,879	0,971
KL	0,890	42,572	0,846	0,968
Li	0,820	43,374	0,803	0,972
Perez-Bueno	0,685	40,054	0,751	0,949
Renyi	0,656	34,075	0,397	0,737
Tosta	0,778	42,036	0,768	0,961
Vahadane	0,749	39,639	0,753	0,948

obter uma tonalidade próxima à imagem de referência. Os resultados com o método da divergência de *Renyi* (Figura 24(f)) mantiveram uma normalização não satisfatória em relação à uniformização dos corantes H&E.

Essa comparação entre diferentes métodos de normalização mostram que eles podem contribuir para melhoramento das cores dos corantes ou afetar as estruturas do tecido representadas. Cada método produziu uma coloração diferente considerando a mesma imagem de referência. Os resultados mostram ainda que a variação da imagem de referência também pode influenciar as informações de brilho da imagem.

Uma avaliação quantitativa foi realizada com os métodos de normalização. As Tabelas 1, 2 e 3 apresentam os valores das métricas com a normalização dos corantes H&E. Quanto mais alto o valor, melhor o resultado proporcionado pela técnica. Os valores marcados em negrito destacam os melhores resultados para cada métrica analisada.

Considerando os resultados com a imagem Ref 1 (Tabela 1), o método Li (Xingyu Li et al. 2015) obteve resultados mais expressivos com a métrica PSNR. Esses resultados são próximos aos valores do método ED. A métrica PSNR permite avaliar as diferenças quantitativas entre a imagem normalizada e original por meio da relação entre a potência máxima que um sinal pode ter e a potência do ruído que distorce sua qualidade (Sara et al. 2019). Ainda há valores próximos entre os resultados das técnicas Li e ED pela

Tabela 2 – Resultados quantitativos das técnicas de normalização pela imagem Ref 2.

Método	FSIM	PSNR	QSSIM	SSIM
ED	0,902	45,170	0,914	0,979
KL	0,896	44,173	0,877	0,974
Li	0,810	42,925	0,753	0,974
Perez-Bueno	0,691	40,703	0,733	0,955
Renyi	0,441	29,808	0,083	0,242
Tosta	0,781	40,820	0,738	0,945
Vahadane	0,734	40,418	0,686	0,954

Tabela 3 – Métricas de normalização dos métodos estudados, a partir da imagem de Ref 3.

Método	FSIM	PSNR	QSSIM	SSIM
ED	0,892	44,791	0,901	0,971
KL	0,867	41,975	0,796	0,956
Li	0,792	39,630	0,707	0,955
Perez-Bueno	0,543	37,020	0,494	0,889
Renyi	0,648	33,575	0,373	0,700
Tosta	0,743	38,850	0,627	0,905
Vahadane	0,757	38,811	0,638	0,930

métrica SSIM. Essa métrica permite avaliar a perda de qualidade da imagem normalizada e a semelhança com a imagem original a partir do cálculo de um valor médio normalizado, o qual indica o nível de similaridade estrutural entre as imagens (Sara et al. 2019). A métrica FSIM representa a similaridade das imagens normalizadas e originais, de modo que seus resultados devem ser maximizados para indicação de bons resultados (Zhang et al. 2011). Neste caso, apenas a ED foi superior às demais técnicas avaliadas. Com a métrica QSSIM, que é uma medida que avalia as degradações na luminância e cromaticidade de uma imagem em comparação à original, os valores mais altos indicam melhores resultados de normalização, com menos degradação da luminância (como borrões e ruídos) e cores mais intensas na imagem (Kolaman et al. 2012). Neste caso, a técnica ED se destacou entre as abordagens investigadas. Nota-se que a abordagem de *Renyi* manteve o desempenho inferior aos demais métodos. Essa informação corrobora as observações qualitativas realizadas pela Figura 22.

Para as imagens Ref 2 e Ref 3 (Tabelas 2 e 3), os resultados seguiram comportamentos diferentes em relação aos modelos com melhor desempenho. Neste caso, o método Li obteve valores mais baixos em relação à Tabela 1. A técnica ED foi superior às demais quando se observam os valores de todas as métricas investigadas. Ressalta-se que a técnica *Renyi* manteve valores baixos quando foram utilizadas outras imagens de referência.

Tabela 4 – Valores das métricas Acurácia e coeficiente *Dice* obtidas na segmentação extraídas na etapa de teste do modelo.

Método	Referência	Acurácia (%)	<i>Dice</i>
Sem normalização	-	86,20	0,7156
ED	Ref 1	86,20	0,7511
	Ref 2	86,07	0,7369
	Ref 3	86,89	0,7523
KL	Ref 1	85,85	0,7291
	Ref 2	86,36	0,7683
	Ref 3	86,00	0,7242
Li	Ref 1	85,85	0,7381
	Ref 2	86,09	0,7019
	Ref 3	86,31	0,7064
Perez-Bueno	Ref 1	85,20	0,6969
	Ref 2	85,88	0,7324
	Ref 3	84,80	0,7331
Renyi	Ref 1	83,67	0,6912
	Ref 2	80,75	0,6126
	Ref 3	83,54	0,6450
Tosta	Ref 1	85,06	0,7140
	Ref 2	85,02	0,6983
	Ref 3	85,39	0,7006
Vahadane	Ref 1	86,65	0,7361
	Ref 2	85,96	0,7210
	Ref 3	85,81	0,6800

5.2 Segmentação de núcleos celulares em imagens histológicas de displasia

Nessa seção, são apresentados os resultados dos experimentos de segmentação utilizando o modelo baseado na U-Net, utilizando os diferentes métodos de normalização. Foram realizados treinamentos considerando as imagens sem normalização e com as estratégias de normalização com diferentes imagens de referência. A arquitetura do modelo foi a mesma para todos os treinamentos. A Tabela 4 apresenta os resultados com as principais métricas de avaliação. Os melhores resultados para cada uma das métricas estão marcados em negrito.

Analisando esses resultados, pode-se notar que a normalização de imagens permitiu melhorar o desempenho do modelo de segmentação dos tecidos da displasia com alguns métodos. Dos 21 modelos resultantes das normalizações, 4 (19% dos modelos) obtiveram acurácias mais altas que o modelo treinado com imagens originais, e 11 (52% dos modelos) obtiveram *Dice* mais altos. Os métodos de normalização ED e KL obtiveram melhores resultados em comparação com as imagens originais e os métodos de normalização. No entanto, isso ocorreu nas diferentes métricas (acurácia e coeficiente *Dice*). A

métrica acurácia traz a proporção total de pixels corretamente rotulados, mas como há um desbalanceamento entre as regiões dos pixels de objetos e a região do fundo, pode apresentar limitações devido a isso. A métrica coeficiente *Dice* é mais robusta para a análise dos modelos, então, parte das discussões foi concentrada nesta métrica.

Nesta análise, a escolha da imagem de referência também influenciou os resultados da segmentação. Dos sete modelos que consideraram normalização, quatro obtiveram melhores métricas *Dice* com a imagem Ref 1, um com a Ref 2 e dois com a Ref 3. A partir dessas quantidades, dá-se a entender que a Ref 1 pode ter sido mais vantajosa nesse processo, apesar do melhor resultado geral ser do uso da Ref 2 pelo método KL. Além disso, os piores resultados com as métricas da segmentação foram obtidos com as imagens normalizadas pelo método *Renyi*, corroborando os resultados obtidos pelas métricas de normalização e tendo seu melhor resultado com a Ref 1.

Na Figura 25 é exemplificada a segmentação da imagem de tecido de displasia. A máscara de segmentação utilizada para validação dos resultados é mostrada na Figura 25(b). Nas Figuras 25(c) e 25(e) são apresentados os resultados da segmentação com a imagem original e normalizada com KL baseada na Ref 2, que obteve o melhor resultado do coeficiente *Dice*, respectivamente. Juntamente com os exemplos de segmentação, são exibidas as imagens que expõem os pixels Falsos Positivos (FP) e Falsos Negativos (FN) quando há comparação da segmentação com a imagem do padrão ouro. Os FP são representados pela cor verde e os FN pela cor vermelha. Nas imagens que exibem os FP e FN, foram adicionadas setas azuis para destacar os núcleos que sofreram mais diferenciação entre os dois modelos de segmentação. Percebe-se que, apesar de visualmente as máscaras parecerem próximas, o uso de normalização de cores afetou o processo de segmentação. Nota-se nas Figuras 25(d) e 25(f) nas setas da cor azul que regiões FN têm variação entre as imagens, original e normalizada. Desse modo, demonstra-se que não é possível validar esses métodos de segmentação visualmente, comprovando a importância do uso das métricas de segmentação.

Nas Figuras 26 e 27 são apresentados os gráficos relacionados às métricas para a etapa de treinamento e validação do modelo de segmentação com as imagens sem normalização e as imagens normalizadas pelo método KL (Ref 2). Esses cenários foram escolhidos devido aos resultados das métricas de avaliação da segmentação. Percebe-se que as redes não apresentaram problemas de *overfitting* e as curvas obtiveram tendências de crescimento. Nota-se uma pequena mudança entre os dados de validação no modelo original (Figura 26) e da arquitetura com as imagens normalizadas (Figura 27), onde as curvas do modelo original estão levemente mais próximas.

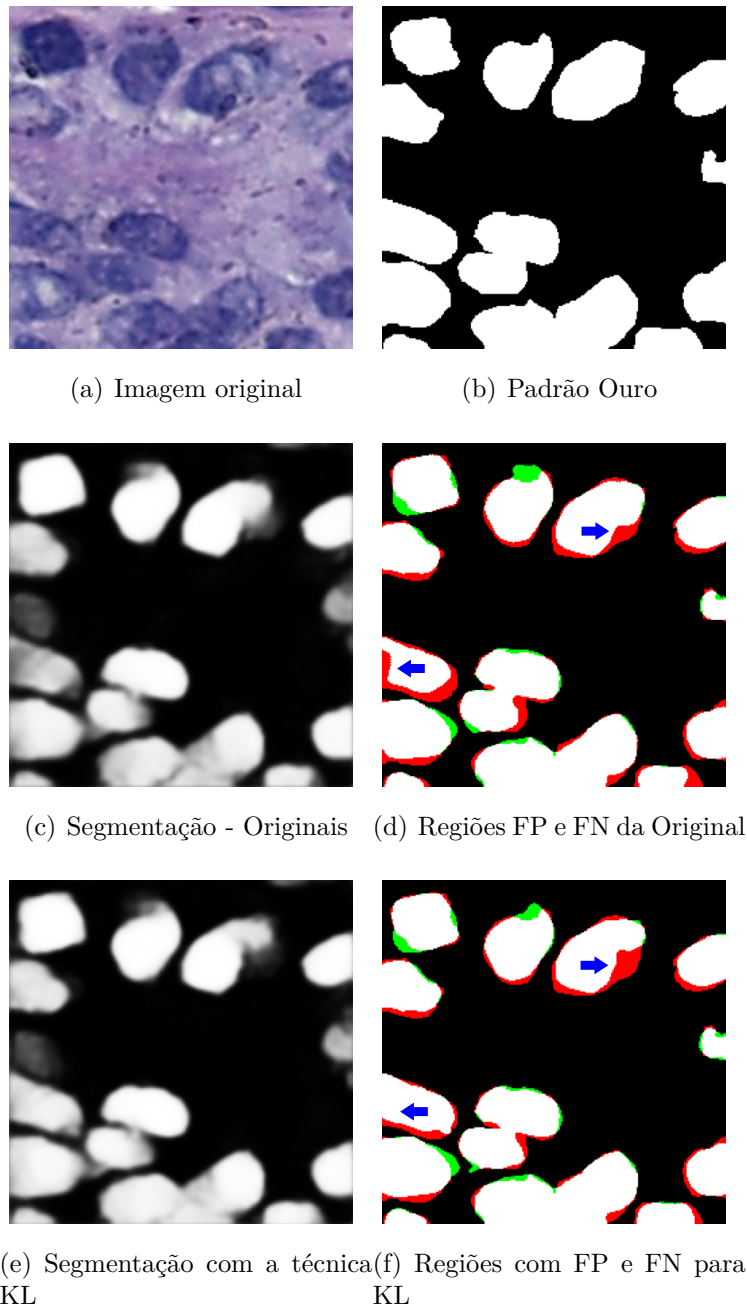


Figura 25 – Exemplos de resultados da segmentação de núcleos celulares, comparando a máscara padrão ouro e os resultados do método KL utilizando a imagem Ref 2. Os FP são representados pela cor verde e os FN pela cor vermelha.

5.3 Interpretação das características na segmentação das imagens

Nesta seção são apresentados os resultados dos experimentos para interpretação das características (*features*) e melhoramento com a técnica Protoseg. Foram analisados o desempenho do parâmetro *SA Score* em cada camada da rede por meio do Protoseg. Esse método ainda possibilitou aprimorar o desempenho do modelo durante o treinamento

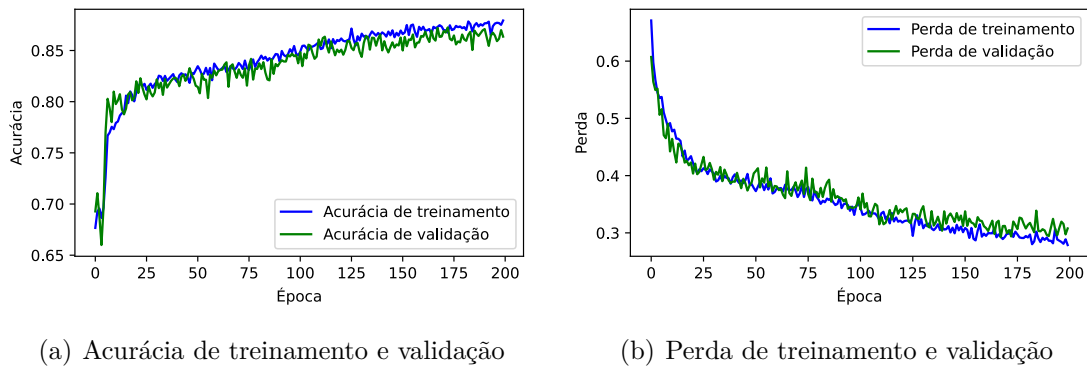


Figura 26 – Gráficos de treinamento do modelo com imagens H&E sem normalização: (esquerda) curva da acurácia versus épocas; (direita) taxa de perda na etapa de treinamento e validação.

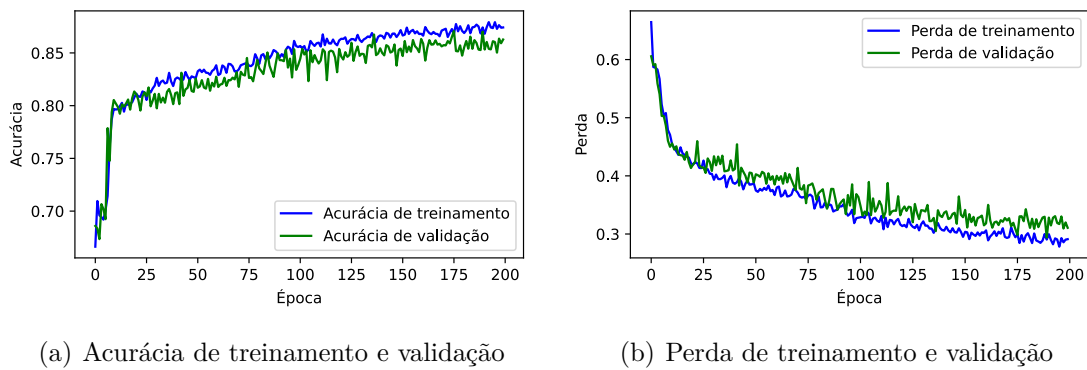


Figura 27 – Gráficos de treinamento da rede com imagens normalizadas com o método KL a partir da imagem Ref 2: (esquerda) curva da acurácia versus épocas; (direita) taxa de perda na etapa de treinamento e validação.

com a abordagem denominada “Protoseg Online”. O modelo original, sem o Protoseg, foi denominado “Protoseg Offline”. Os gráficos dos modelos são apresentados nas Figuras 28 e 29 em que mostram os valores médios e desvios calculados considerando as camadas do modelo de segmentação para o online (curva na cor verde) e offline (curva na cor azul), respectivamente.

As pontuações *SA Score* em todos os blocos (camadas) do modelo de segmentação demonstram a evolução da métrica e melhoramento da segmentação entre as duas curvas (offline e online) (Figura 28). Esse comportamento de melhoramento ocorreu em todas as camadas convolucionais do modelo da imagem sem normalização, mostrando que a abordagem Protoseg contribuiu para melhorar o desempenho do modelo. Quando o procedimento foi empregado sobre o modelo com as imagens normalizadas com a técnica KL a partir da imagem Ref 2 (Figura 29), observou-se que os valores *SA Score* ficaram próximos entre estratégias (online e offline). Isso mostra que os valores offline apresentam desempenho melhor com a normalização e o uso do melhoramento com o ProtoSeg não permitiu uma evolução do melhoramento do modelo. Isso também é notado na variação

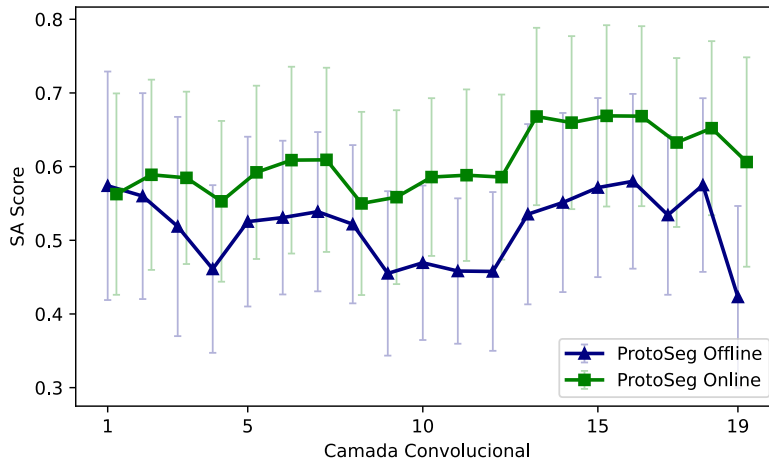


Figura 28 – Valores médios e desvios padrões dos *SA Scores* de cada camada do modelo com imagens não normalizadas com e sem uso da estratégia online do ProtoSeg.

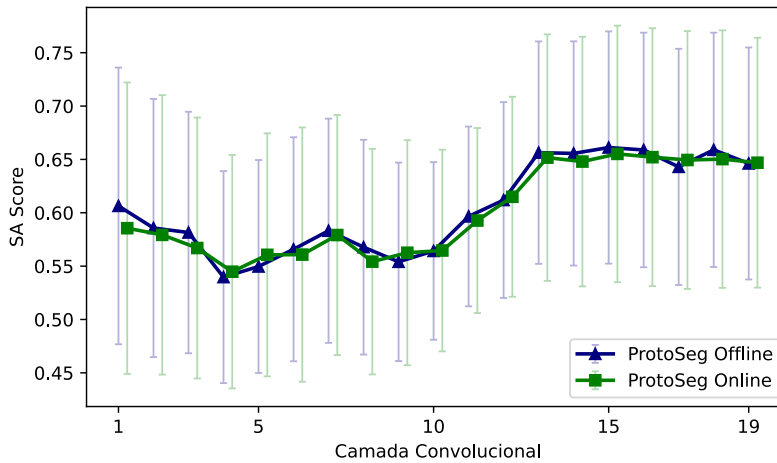


Figura 29 – Resultados com as médias e desvios padrões dos *SA Scores* calculados para cada camadas da rede treinada com imagens normalizadas pelo método KL a partir da imagem Ref 2.

dos desvios-padrão em cada camada do modelo.

A Tabela 5 mostra os resultados dos experimentos com o modelo de rede proposto com a aplicação da abordagem ProtoSeg. O modelo foi treinado com as imagens obtidas por cada técnica de normalização e originais.

Numa análise das imagens originais (sem a normalização) nota-se que a abordagem permitiu melhorar o desempenho do modelo, pois os valores do coeficiente *Dice* foram de 0,7156 para 0,7633, o que representa uma elevação de 6,66%. No caso da métrica acurácia, esse comportamento de variação foi pequeno. No entanto, como já destacado anteriormente, essa métrica apresenta algumas restrições para uma análise entre os pixels da região de fundo e objeto. Quando são observados os resultados apenas das abordagens

Tabela 5 – Resultados com as métricas de segmentação com o método de segmentação e a abordagem Protoseg.

Método	Referência	Acurácia (%)	Dice
Sem normalização	-	86,71	0,7633
ED	Ref 1	86,27	0,7393
	Ref 2	86,06	0,7561
	Ref 3	85,80	0,7376
KL	Ref 1	86,24	0,7491
	Ref 2	86,05	0,7573
	Ref 3	86,39	0,7216
Li	Ref 1	85,97	0,7405
	Ref 2	86,27	0,7155
	Ref 3	86,17	0,7454
Perez-Bueno	Ref 1	84,91	0,6939
	Ref 2	85,34	0,7187
	Ref 3	85,44	0,7344
Renyi	Ref 1	82,42	0,6855
	Ref 2	81,92	0,6840
	Ref 3	82,98	0,6624
Tosta	Ref 1	85,40	0,6833
	Ref 2	85,42	0,7212
	Ref 3	84,98	0,7263
Vahadane	Ref 1	85,93	0,7117
	Ref 2	86,04	0,7454
	Ref 3	86,37	0,7360

com os métodos de normalização, observa-se que o método KL foi o que apresentou os melhores valores para a métrica coeficiente *Dice* com a imagem Ref 2. No entanto, esse valor não foi superior ao resultado obtido com o modelo sem uso do ProtoSeg (Tabela 4 (coeficiente *Dice* 0,7683)). Desse modo, percebe-se que o método Protoseg, com abordagem online, em que ocorre o ajuste baseado na função *SA score*, não contribuiu para elevar os valores das métricas quando aplicado sobre as imagens normalizadas. O modelo que obteve os resultados mais baixos foi novamente o que realizou o treinamento com as imagens normalizadas pelo método *Renyi*, mesmo com Protoseg.

Os gráficos de treinamento das redes sem normalização e com normalização utilizando o método KL estão expostos nas Figuras 30 e 31. Os gráficos exibem as curvas de acurácia e perda de treinamento e validação em que é possível notar uma maior variação comparada com os modelos que não utilizaram o método Protoseg (Figuras 26 e 27). Isso é uma característica negativa, pois uma maior variação indica que o treinamento passou por um processo mais instável, podendo afetar negativamente o resultado final do modelo.

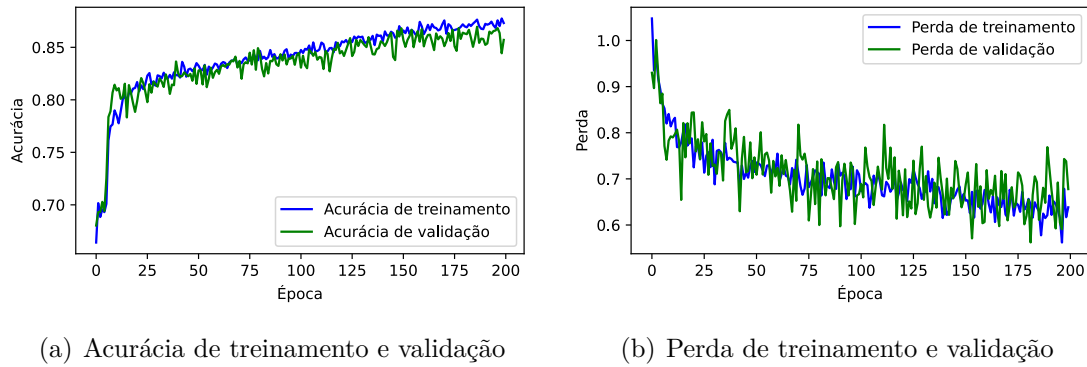


Figura 30 – Gráficos de treinamento da rede treinada com imagens sem normalização com o método Protoseg: (esquerda) curva da acurácia versus épocas; (direita) taxa de perda na etapa de treinamento e validação.

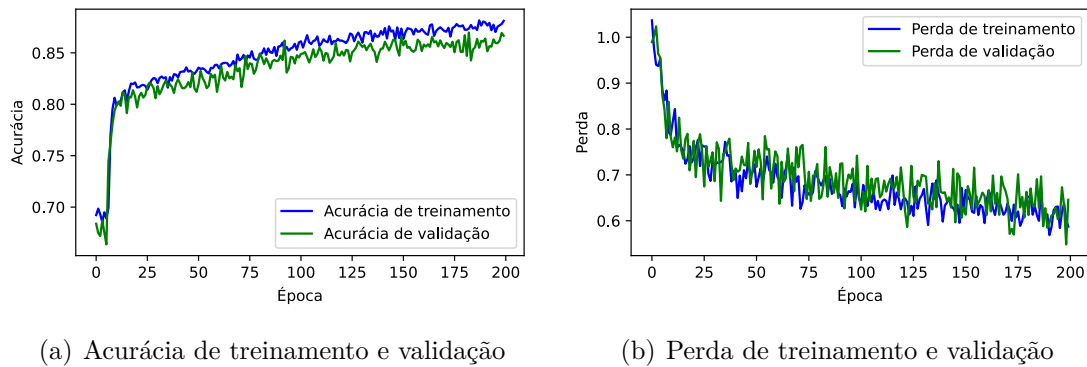


Figura 31 – Gráficos de treinamento da rede treinada com imagens normalizadas com o método KL a partir da imagem Ref 2 e com o método Protoseg Online: (esquerda) curva da acurácia versus épocas; (direita) taxa de perda na etapa de treinamento e validação.

5.4 Exploração do Conjunto de Dados com os Modelos de Aprendizagem Profunda

Esta seção trata de uma avaliação comparativa entre as bases de imagens para análise de generalização e avaliação do impacto da normalização. Foram realizados experimentos entre os modelos treinados com as imagens originais (sem normalização) e o modelo treinado com as imagens normalizadas pelo método KL. Esse modelo foi selecionado devido aos resultados mais promissores nas etapas anteriores.

O modelo denominado *A* representa a rede de segmentação treinada com as imagens originais e avaliada na etapa de teste as imagens normalizadas pelo método KL (Ref 2). O modelo *B* representa a rede de segmentação treinada com as imagens originais e avaliada (teste) com as imagens normalizadas com o método KL. O modelo *C* representa a rede treinada com as imagens normalizadas com o método KL sem o Protoseg e testada com as imagens originais. O modelo *D* representa o modelo treinado com os dados normalizados

Tabela 6 – Avaliação da generalização do modelo e influência da normalização sobre os dados.

Modelo	Acurácia	<i>Dice</i>
A	74,62%	0,2029
B	78,29%	0,4093
C	77,52%	0,3658
D	77,90%	0,3902

Tabela 7 – Avaliação entre as diferentes abordagens para segmentação de núcleos em imagens histológicas de displasia.

Estudos	Normalização	Acurácia	<i>Dice</i>
(D. F. D. d. Santos et al. 2020a)	-	87,90%	0,8200
(A. B. Silva et al. 2022)	-	89,31%	0,8200
(A. B. Silva et al. 2023)	-	91,95%	0,8700
(A. B. Silva et al. 2024b)	Reinhard	92,38%	0,8300
Método Proposto	-	86,20%	0,7156
Método Proposto com Protoseg	-	86,71%	0,7633
Método Proposto	KL	86,36%	0,7683
Método Proposto com Protoseg	KL	86,05%	0,7573

pelo método KL e Protoseg Online, e testado com as imagens originais. Na Tabela 6 são apresentados os resultados das avaliações realizadas entre os modelos investigados.

A partir dos resultados na Tabela 6, percebe-se que a avaliação de generalização mostra o impacto da normalização. Os modelos treinados a partir das imagens originais (*A* e *C*) geraram valores para as métricas *Dice* muito baixos quando testados nas imagens normalizadas pelo método KL. Da mesma maneira, os modelos treinados a partir das imagens normalizadas pelo método KL (*B* e *D*) obtiveram resultados baixos quando testados com as imagens sem normalização. Dessa forma, nota-se que a normalização afeta a segmentação de núcleos celulares em imagens histológicas com uso do modelo de segmentação. Isso mostra que as imagens de teste dos modelos devem seguir o mesmo padrão de cores que foram usadas no modelo durante a etapa de treinamento.

5.5 Avaliação com o Estado da Arte

Nesta seção são apresentados estudos que exploraram a base de imagens investigada para segmentação de núcleos. Os resultados são mostrados na Tabela 7 em relação às métricas acurácia e coeficiente *Dice* para os modelos com e sem o Protoseg.

O trabalho de (A. B. Silva et al. 2024b) é um dos primeiros estudos que exploraram a variação de imagens de referência com a normalização na tarefa de segmentação. Esse modelo investigou uma rede semântica baseada na U-net com um *backbone* da ResNet para segmentação com diferentes técnicas de normalização. Os resultados mostram que a normalização influenciou os valores das métricas durante o processo de segmentação. No

caso desse estudo, a abordagem de normalização contribuiu para melhorar o modelo de segmentação. Em relação a este trabalho, os resultados foram menores, mas é importante destacar que os métodos empregados para normalização neste estudo são diferentes do trabalho de (A. B. Silva et al. 2024b). Essa pesquisa investigou métodos de normalização baseados em algoritmos de aprendizado de dicionário. Isso pode ser um dos motivos da diferença em relação aos resultados do trabalho de (A. B. Silva et al. 2024b). Um outro ponto ainda a ser considerado é que os valores sem normalização são próximos aos resultados obtidos no estudo proposto nesta pesquisa (acurácias de 87,84 e 86,71).

Em relação ao estudo de (A. B. Silva et al. 2023), os autores propuseram uma estratégia *ensemble* para segmentação com modelos semânticos para o conjunto de dados. Neste caso, o modelo traz resultados mais relevantes. Ressalta-se que essa abordagem explora a composição de comitês com mais de uma arquitetura CNN para avaliação da segmentação das estruturas, o que pode ser mais efetivo na segmentação. Outros métodos como (D. F. D. d. Santos et al. 2020a) e (A. B. Silva et al. 2022), tiveram resultados próximos pela métrica de acurácia. Destaca-se que o modelo de rede em (A. B. Silva et al. 2022) foi uma rede Mask R-CNN, a qual é uma abordagem por instância, o que pode apresentar resultados superiores em relação aos modelos semânticos. Os dados de (D. F. D. d. Santos et al. 2020a) são próximos em relação à métrica acurácia. No entanto, esse método tem resultados superiores em relação ao coeficiente *Dice*. Apesar das limitações, o estudo atual permite analisar a interpretação das características das camadas convolucionais do modelo, assim como a influência de modelos de normalização sobre as estruturas da CNN.

Conclusão

Neste trabalho foi desenvolvido um estudo da combinação das etapas de normalização de corantes e segmentação com interpretação dos descritores aprendidos. O estudo avaliou sete métodos de normalização em imagens histológicas coradas com H&E a partir de três imagens de referência diferentes. Também foi realizada a análise da influência das imagens resultantes de cada método em um modelo CNN treinado e testado utilizando essas imagens, com o intuito de realizar a segmentação de núcleos celulares presentes nesses tecidos. No modelo de segmentação foi considerado o método de interpretação dos descritores de uma CNN com ProtoSeg, o que pode quantificar as capacidades de discriminação em diferentes camadas da rede. Essa abordagem permite fornecer uma métrica de pontuação média, oferecendo uma estimativa de desempenho da saída nas imagens de teste com possível aprimoramento de desempenho da rede com ajuste adicionado na função de perda.

Ao comparar os métodos de normalização ED, KL, Li, Perez-Bueno, Renyi, Tosta e Vahadane, a partir dos resultados das métricas FSIM, PSNR, QSSIM e SSIM, foi verificado que os valores variaram, porém em intervalo pequeno. Entre eles, os métodos ED e Li foram os que mais se destacaram, com os valores mais altos nas métricas, como visto na Tabela 1, com FSIM de 0,972 para o Li, na Tabela 2 com FSIM de 0,979 para o ED e na Tabela 3 com FSIM de 0,971 para o ED. Com isso, é possível inferir que esses métodos de normalização foram os que melhor conseguiram transferir as cores presentes nas imagens de referência para as imagens normalizadas, ainda mantendo as propriedades estruturais das imagens originais.

Com o conjunto de dados originais e os métodos de normalização e imagens de referências, foram obtidas 21 (vinte e um) bases de imagens contendo 456 imagens cada, com resolução de 450×250 pixels. Com o processo de *cropping* e aumento de dados utilizando o método RCAug, o modelo de segmentação foi avaliado com 2.552 imagens para treinamento, 90 imagens para validação e 184 para teste. A rede de segmentação foi avaliada sobre esse mesmo banco de imagens normalizado com cada um dos diferentes métodos analisados. Os resultados das métricas de Acurácia e Coeficiente *Dice* desses modelos

foram apresentados na Tabela 4. O modelo treinado com as imagens normalizadas pelo método KL (considerando a imagem de referência 2) obteve os melhores valores, visto que a métrica *Dice* é mais relevante para a avaliação dos resultados de segmentação. Esse resultado implica que os métodos de normalização de cores podem afetar positivamente o processo de segmentação de imagens, visto que os modelos treinados com as imagens normalizadas pelos métodos ED, KL e Li foram capazes de gerar resultados de segmentação melhores do que o modelo que não considerou a normalização de cores (imagens originais).

O estudo ainda traz uma sequência de treinamentos semelhantes, mas considerando a métrica *SA Score* na função de perda dos modelos. Essa abordagem é denominada Protoseg Online, e os resultados dos testes dos modelos obtidos foram apresentados na Tabela 5. Nesse caso, a metodologia que obteve as melhores métricas foi a que não considerou a normalização de cores, com acurácia de 86,71% e Coeficiente *Dice* de 0,7633. Esses valores são maiores que os obtidos pelo modelo Protoseg Offline sem normalização, de acurácia 86,20% e *Dice* de 0,7156. Desse modo, o Protoseg Online obteve resultados melhores quando não considerou a normalização de cores, e esses resultados também foram melhores que o Protoseg Offline sem normalização de cores. Portanto, pode-se afirmar que a combinação entre Protoseg Online e normalização de cores não é relevante para o domínio de lesões histológicas de displasia, mas a utilização desses métodos de forma individual pode ser benéfica para modelos de segmentação de núcleos celulares.

Por fim, uma análise explicativa das redes foi realizada a partir da plotagem da métrica *SA Score* em cada camada das redes, para as metodologias que consideraram as imagens originais (sem normalização) e as imagens normalizadas pelo método KL a partir da imagem Ref 2, que obtiveram os melhores resultados dos modelos de Protoseg Online e Offline. Nos gráficos dos modelos sem normalização de cores (representação dos descritores em cada camada do modelo), observa-se que a arquitetura que considerou o ProtoSeg Online obteve valores mais altos na maioria das camadas, significando um aprimoramento em relação ao outro modelo. Na análise do gráfico dos modelos que consideraram normalização dos corantes, denota-se que as curvas de desempenho nas camadas da rede (descritores) dos modelos foram de comportamento próximo, indicando que a utilização da metodologia ProtoSeg Online não causou avanços nesse teste. Esses gráficos se correlacionam com a análise realizada sobre as métricas de teste dos modelos, pois esses resultados reafirmam que a utilização do método ProtoSeg só foi produtiva quando não considerada a normalização de cores.

6.1 Principais Contribuições

O estudo desenvolvido mostrou que a utilização de técnicas de normalização de cores dos corantes H&E baseadas em aprendizado de dicionário proporciona melhorias signifi-

cativas nos resultados de métodos de segmentação, em comparação com o uso de imagens não normalizadas, como mostrado nos resultados dos modelos de ProtoSeg Offline. A incorporação de métodos de interpretação no processo de segmentação, como no caso do ProtoSeg Online, permite o desenvolvimento de redes neurais com desempenho superior em relação às métricas analisadas, mas somente quando não se considera o processo de normalização.

A metodologia apresentada neste trabalho pode ser utilizada como ferramenta de estudos para trabalhos futuros. O código-fonte deste trabalho está disponível no GitHub, <https://github.com/AndreDF98/ProtoSeg>.

6.2 Trabalhos Futuros

Para trabalhos futuros, podem ser abordados outros métodos de segmentação de imagens, como modelos baseados em segmentação por instância e modelos *ensemble*, assim como a utilização desses métodos em outras bases de imagens, para avaliar a generalização desses resultados. Além disso, pode ser abordado o uso dos métodos de normalização de cores em etapas de classificação das imagens, a fim de analisar se sua utilização é realmente válida para obter melhores desempenhos. Também é importante avaliar e incorporar novas abordagens de interpretabilidade das redes neurais para melhoramento dos métodos empregados na segmentação, visto que seu uso neste estudo foi promissor em alguns cenários.

6.3 Contribuições em Produção Bibliográfica

- Tosta, T. A. A., Freitas, A. D., de Faria, P. R., Neves, L. A., Martins, A. S., & do Nascimento, M. Z. (2023). A stain color normalization with robust dictionary learning for breast cancer histological images processing. *Biomedical Signal Processing and Control*, 85, 104978.
- Silva, A. F. Q., Freitas, A. D., de Faria, P. R., Neves, L. A., do Nascimento, M. Z., Tosta, T. A. A. (2025) Color normalization by dictionary learning with nuclear segmentation evaluation in H&E histological images aceito para o IEEE CBMS 2025.

Bibliografia

- Ait Skourt, Brahim, Abdelhamid El Hassani e Aicha Majda (2018). “Lung CT Image Segmentation Using Deep Neural Networks”. Em: *Procedia Computer Science* 127. PROCEEDINGS OF THE FIRST INTERNATIONAL CONFERENCE ON INTELLIGENT COMPUTING IN DATA SCIENCES, ICDS2017, pp. 109–113. ISSN: 1877-0509.
- Asgari Taghanaki, Saeid, Kumar Abhishek, Joseph Paul Cohen, Julien Cohen-Adad e Ghassan Hamarneh (jan. de 2021). “Deep semantic segmentation of natural and medical images: a review”. Em: *Artificial Intelligence Review* 54.1, pp. 137–178. ISSN: 1573-7462. DOI: <10.1007/s10462-020-09854-1>. URL: <<https://doi.org/10.1007/s10462-020-09854-1>>.
- Basu, Anusua, Pradip Senapati, Mainak Deb, Rebika Rai e Krishna Gopal Dhal (fev. de 2024). “A survey on recent trends in deep learning for nucleus segmentation from histopathology images”. Em: *Evolving Systems* 15.1, pp. 203–248. ISSN: 1868-6486. DOI: <10.1007/s12530-023-09491-3>. URL: <<https://doi.org/10.1007/s12530-023-09491-3>>.
- Belsare, AD, MM Mushrif, MA Pangarkar e N Meshram (2016). “Breast histopathology image segmentation using spatio-colour-texture based graph partition method”. Em: *Journal of microscopy* 262.3, pp. 260–273.
- Brennan, Patrick C, Aarthi Ganesan, Miguel P Eckstein, Ernest Usang Ekpo, Kriscia Tapia, Claudia Mello-Thoms, Sarah Lewis e Mordechai Z Juni (2019). “Benefits of independent double reading in digital mammography: a theoretical evaluation of all possible pairing methodologies”. Em: *Academic radiology* 26.6, pp. 717–723.
- Chan, Heang-Ping, Lubomir M Hadjiiski e Ravi K Samala (2020a). “Computer-aided diagnosis in the era of deep learning”. Em: *Medical physics* 47.5, e218–e227.
- (2020b). “Computer-aided diagnosis in the era of deep learning”. Em: *Medical Physics* 47.5, e218–e227. DOI: <<https://doi.org/10.1002/mp.13764>>.
- Chen, Chiao-Min, Yao-Sian Huang, Pei-Wei Fang, Cher-Wei Liang e Ruey-Feng Chang (2020). “A computer-aided diagnosis system for differentiation and delineation of ma-

- lignant regions on whole-slide prostate histopathology image using spatial statistics and multidimensional DenseNet”. Em: *Medical physics* 47.3, pp. 1021–1033.
- El-Dahshan, El-Sayed A, Heba M Mohsen, Kenneth Revett e Abdel-Badeeh M Salem (2014). “Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm”. Em: *Expert systems with Applications* 41.11, pp. 5526–5545.
- Devarajan, Karthik e Guoli Wang (2016). “hpcNMF: A high-performance toolbox for non-negative matrix factorization”. Em.
- Devarajan, Karthik, Guoli Wang e Nader Ebrahimi (2015). “A unified statistical approach to non-negative matrix factorization and probabilistic latent semantic indexing”. Em: *Machine Learning* 99.1, pp. 137–163.
- Dhaini, Mohamad, Maxime Berar, Paul Honeine e Antonin Van Exem (2023). “Unsupervised domain adaptation for regression using dictionary learning”. Em: *Knowledge-Based Systems* 267, p. 110439.
- Dost, F., K. Lê Cao, P.J. Ford, C. Ades e C.S. Farah (2014). “Malignant transformation of oral epithelial dysplasia: a real-world evaluation of histopathologic grading”. Em: *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology* 117.3, pp. 343–352. ISSN: 2212-4403.
- Fonseca-Silva, Thiago, Marina G Diniz, Sílvia F de Sousa, Ricardo S Gomez e Carolina C Gomes (2016). “Association between histopathological features of dysplasia in oral leukoplakia and loss of heterozygosity”. Em: *Histopathology* 68.3, pp. 456–460.
- Gal, Yarin e Zoubin Ghahramani (2016). *Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning*. arXiv: <1506.02142> [stat.ML]. URL: <<https://arxiv.org/abs/1506.02142>>.
- Goodfellow, Ian, Yoshua Bengio e Aaron Courville (2016). *Deep Learning*. <<http://www.deeplearningbook.org>>. MIT Press.
- Gourdeau, Daniel, Simon Duchesne e Louis Archambault (2022). “On the proper use of structural similarity for the robust evaluation of medical image synthesis models”. Em: *Medical Physics* 49.4, pp. 2462–2474.
- Hameed, Zabit, Begonya Garcia-Zapirain, José Javier Aguirre e Mario Arturo Isaza-Ruget (2022). “Multiclass classification of breast cancer histopathology images using multilevel features of deep convolutional neural network”. Em: *Scientific Reports* 12.1, p. 15600.
- Hao, Shijie, Yuan Zhou e Yanrong Guo (2020). “A Brief Survey on Semantic Segmentation with Deep Learning”. Em: *Neurocomputing* 406, pp. 302–321.
- Haykin, Simon (2009). *Neural networks and learning machines*. Pearson education Upper Saddle River.
- He, Sheng, Yanfang Feng, P. Ellen Grant e Yangming Ou (2023). “Segmentation ability map: Interpret deep features for medical image segmentation”. Em: *Medical Image Analysis* 84.

- He, Wenbin, Yongjie Han, Wuyi Ming, Jinguang Du, Yinxia Liu, Yuan Yang, Leijie Wang, Yongqiang Wang, Zhiwen Jiang, Chen Cao e Jie Yuan (2022). “Progress of Machine Vision in the Detection of Cancer Cells in Histopathology”. Em: *IEEE Access* 10, pp. 46753–46771.
- Humeau-Heurtier, Anne (2019). “Texture feature extraction methods: A survey”. Em: *IEEE Access* 7, pp. 8975–9000.
- Janowczyk, Andrew e Anant Madabhushi (2016). “Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases”. Em: *Journal of Pathology Informatics* 7.1, p. 29. ISSN: 2153-3539. DOI: <<https://doi.org/10.4103/2153-3539.186902>>. URL: <<https://www.sciencedirect.com/science/article/pii/S2153353922005478>>.
- Junior, Wagner Lopes Moreira (2018). “Uma Nova Abordagem de Descritor de Textura Baseada em Transformada Ripplet para Classificação de Lesões da Mama”. Diss. de mest. Universidade Federal do ABC.
- Kademan, Deepak (2019). *Improving Outcomes in Oral Cancer: A Clinical and Translational Update*. Springer Nature.
- Kang, H., D. Luo, W. Feng, S. Zeng, T. Quan, J. Hu e X. Liu (2021). “Stainnet: a fast and robust stain normalization network”. Em: *Frontiers in Medicine*.
- Kolaman, Amir e Orly Yadid-Pecht (2012). “Quaternion Structural Similarity: A New Quality Index for Color Images”. Em: *IEEE Transactions on Image Processing* 21.4, pp. 1526–1536.
- Krizhevsky, Alex, Ilya Sutskever e Geoffrey E. Hinton (2017). “ImageNet Classification with Deep Convolutional Neural Networks”. Em: 60.6, pp. 84–90.
- Kumar, Neeraj, Ruchika Verma, Deepak Anand, Yanning Zhou, Omer Fahri Onder, Efstathios Tsougenis, Hao Chen, Pheng-Ann Heng, Jiahui Li, Zhiqiang Hu et al. (2019). “A multi-organ nucleus segmentation challenge”. Em: *IEEE transactions on medical imaging* 39.5, pp. 1380–1391.
- Kumar, Vinay, Abul K Abbas e Jon C Aster (2013). *Robbins Patologia Básica*. Elsevier Brasil.
- Kumar, Vinay, Jon C Aster e Abbas Abbas (2010). *Robbins and Cotran Patologia-Bases Patológicas das Doenças*. Elsevier Brasil.
- Lecun, Y., L. Bottou, Y. Bengio e P. Haffner (1998). “Gradient-based learning applied to document recognition”. Em: *Proceedings of the IEEE* 86.11, pp. 2278–2324.
- Lei, G., Y. Xia, D.H. Zhai, W. Zhang, D. Chen e D. Wang (2020). “Staincnns: An efficient stain feature learning method”. Em: *Neurocomputing*, pp. 267–273.
- Li, Xingyu e Konstantinos N. Plataniotis (2015). “A Complete Color Normalization Approach to Histo-pathology Images Using Color Cues Computed From Saturation-Weighted Statistics”. Em: *IEEE Transactions on Biomedical Engineering*.

- Li, Xintong, Chen Li, Md Mamunur Rahaman, Hongzan Sun, Xiaoqi Li, Jian Wu, Yudong Yao e Marcin Grzegorzec (ago. de 2022). “A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches”. Em: *Artificial Intelligence Review* 55.6, pp. 4809–4878. ISSN: 1573-7462. DOI: <10.1007/s10462-021-10121-0>. URL: <<https://doi.org/10.1007/s10462-021-10121-0>>.
- Liu, Menglu, Junyu Dong, Xinghui Dong, Hui Yu e Lin Qi (2018). “Segmentation of lung nodule in CT images based on mask R-CNN”. Em: *2018 9th International Conference on Awareness Science and Technology (iCAST)*. IEEE, pp. 1–6.
- Macenko, M., M. Niethammer, J.S. Marron, D. Borland, J.T. Woosley, Xiaojun Guan, C. Schmitt e N.E. Thomas (2009). “A method for normalizing histology slides for quantitative analysis”. Em: *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, pp. 1107–1110.
- Mairal, J., F. Bach, J. Ponce, G. Sapiro, R. Jenatton e G. Obozinski (2017). *SPArse Modeling Software*. <<https://bit.ly/2Klt3fs>>, Acessado em: 14/Nov/2019.
- Mairal, Julien, Francis Bach, Jean Ponce e Guillermo Sapiro (2010). “Online learning for matrix factorization and sparse coding”. Em: *Journal of Machine Learning Research* 11, pp. 19–60.
- Minaee, Shervin, Yuri Y Boykov, Fatih Porikli, Antonio J Plaza, Nasser Kehtarnavaz e Demetri Terzopoulos (2021). “Image segmentation using deep learning: A survey”. Em: *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Minh, Dang, H. Xiang Wang, Y. Fen Li e Tan N. Nguyen (jun. de 2022). “Explainable artificial intelligence: a comprehensive review”. Em: *Artificial Intelligence Review* 55.5, pp. 3503–3568. ISSN: 1573-7462.
- Molnar, Christoph (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. <<https://christophm.github.io/interpretable-ml-book/>>.
- Müller, Susan (2018). “Oral epithelial dysplasia, atypical verrucous lesions and oral potentially malignant disorders: focus on histopathology”. Em: *Oral surgery, oral medicine, oral pathology and oral radiology* 125.6, pp. 591–602.
- Nasir, Esha Sadia, Arshi Parvaiz e Muhammad Moazam Fraz (ago. de 2023). “Nuclei and glands instance segmentation in histology images: a narrative review”. Em: *Artificial Intelligence Review* 56.8, pp. 7909–7964. ISSN: 1573-7462. DOI: <10.1007/s10462-022-10372-5>. URL: <<https://doi.org/10.1007/s10462-022-10372-5>>.
- Oliveira Santos, Marceli de (2020). “Estimativa/2020 Incidência de Câncer no Brasil”. Em: *Revista Brasileira de Cancerologia* 66.1.
- Organização Mundial da Saúde (2020). *WHO report on cancer: setting priorities, investing wisely and providing care for all*. World Health Organization.
- Pérez-Bueno, Fernando, Juan G. Serra, Miguel Vega, Javier Mateos, Rafael Molina e Aggelos K. Katsaggelos (2022). “Bayesian K-SVD for H and E blind color deconvolution.

- Applications to stain normalization, data augmentation and cancer classification”. Em: *Computerized Medical Imaging and Graphics* 97.
- Rabinovich, Andrew, Sameer Agarwal, Casey Laris, Jeffrey Price e Serge Belongie (2003). “Unsupervised Color Decomposition Of Histologically Stained Tissue Samples”. Em: *Advances in Neural Information Processing Systems*. Ed. por S. Thrun, L. Saul e B. Schölkopf. Vol. 16. MIT Press.
- Ramadan, Saleem Z. (2020). “Methods Used in Computer-Aided Diagnosis for Breast Cancer Detection Using Mammograms: A Review”. Em: *Hindawi - Journal of Healthcare Engineering*.
- Reinhard, E., M. Adhikhmin, B. Gooch e P. Shirley (2001). “Color transfer between images”. Em: *IEEE Computer graphics and applications*, pp. 34–41.
- Rong, Ruichen, Shidan Wang, Xinyi Zhang, Zhuoyu Wen, Xian Cheng, Liwei Jia, Donghan M Yang, Yang Xie, Xiaowei Zhan e Guanghua Xiao (2023). “Enhanced Pathology Image Quality with Restore-Generative Adversarial Network”. Em: *The American Journal of Pathology* 193.4, pp. 404–416.
- Ronneberger, O., P.Fischer e T. Brox (2015). “U-Net: Convolutional Networks for Biomedical Image Segmentation”. Em: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Vol. 9351. LNCS. (available on arXiv:1505.04597 [cs.CV]). Springer, pp. 234–241. URL: <<http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a>>.
- Rosebrock, Dr. Adrian (2017). *Deep Learning for Computer Vision with Python*. PyImageSearch.
- Runz, M., D. Rusche, S. Schmidt, M.R. Weihrauch, J. Hesser e C.A. Weis (2021). “Normalization of he-stained histological images using cycle consistent generative adversarial networks”. Em: *Diagnostic Pathology*, pp. 1–10.
- Salehi, P. e A. Chalechale (2020). “Pix2pix-based stain-to-stain translation: A solution for robust stain normalization in histopathology images analysis”. Em: *2020 International Conference on Machine Vision and Image Processing (MVIP)*, pp. 1–7.
- Santos, Dalí Freire Dias dos (2022). “Automated segmentation of tumor regions from oral histological whole slide images using fully convolutional neural networks”. Tese de dout. Universidade Federal de Uberlândia.
- Santos, Dalí Freire Dias dos, Adriano Barbosa Silva, Paulo de Faria, Bruno Travençolo e Marcelo do Nascimento (2020a). “Impacts of color space transformations on dysplastic nuclei segmentation using CNN”. Em: *Workshop de Visão Computacional (WVC)*. SBC, pp. 6–11.
- Santos, Dalí Freire Dias dos, Thaína A. A. Tosta, Adriano Barbosa Silva, Paulo R. de Faria, Bruno A. N. Travençolo e Marcelo Z. do Nascimento (2020b). “Automated Nuclei Segmentation on Dysplastic Oral Tissues Using CNN”. Em: *2020 International*

- Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 45–50. DOI: <10.1109/IWSSIP48289.2020.9145157>.
- Santos, Marceli de Oliveira, Fernanda Cristina da Silva de Lima, Luís Felipe Leite Martins, Julio Fernando Pinto Oliveira, Liz Maria de Almeida e Marianna de Camargo Cancela (fev. de 2023). “Estimativa de Incidência de Câncer no Brasil, 2023-2025”. Em: *Revista Brasileira de Cancerologia* 69.1, e–213700.
- Santos Pereira, Joabe dos, Marianne de Vasconcelos Carvalho, Águida Cristina Gomes Henriques, Tiago Henrique de Queiroz Camara, Márcia Cristina da Costa Miguel e Roseana de Almeida Freitas (2011). “Epidemiology and correlation of the clinico-pathological features in oral epithelial dysplasia: analysis of 173 cases”. Em: *Annals of Diagnostic Pathology* 15.2, pp. 98–102. ISSN: 1092-9134.
- Sara, Umme, Morium Akter e Mohammad Shorif Uddin (2019). “Image Quality Assessment through FSIM, SSIM, MSE and PSNRA Comparative Study”. Em: *Journal of Computer and Communications*. URL: <<https://api.semanticscholar.org/CorpusID:104425037>>.
- Setiadi, De Rosal Igantius Moses (mar. de 2021). “PSNR vs SSIM: imperceptibility quality assessment for image steganography”. Em: *Multimedia Tools and Applications* 80.6, pp. 8423–8444. ISSN: 1573-7721.
- Sha, L., D. Schonfeld e A. Sethi (2017). “Color normalization of histology slides using graph regularized sparse nmf”. Em: *Medical Imaging*.
- Shaban, M.T., C. Baur, N. Navab e S. Albarqouni (2019). “Staingan: Stain style transfer for digital histological images”. Em: *IEEE 16th international symposium on biomedical imaging*, pp. 953–956.
- Silva, Adriano Barbosa (2019). “Métodos Computacionais para Análise e Classificação de Displasias em Imagens da Cavidade Bucal”. Diss. de mestr. Universidade Federal de Uberlândia - Faculdade de Computação.
- Silva, Adriano Barbosa, Alessandro S. Martins, Leandro A. Neves, Paulo R. Faria, Thaína A. A. Tosta e Marcelo Zanchetta do Nascimento (2019). “Automated Nuclei Segmentation in Dysplastic Histopathological Oral Tissues Using Deep Neural Networks”. Em: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. Ed. por Ingela Nyström, Yanio Hernández Heredia e Vladimir Milián Núñez. Cham: Springer International Publishing, pp. 365–374. ISBN: 978-3-030-33904-3.
- Silva, Adriano Barbosa, Alessandro Santana Martins, Thaína Aparecida Azevedo Tosta, Adriano Mota Loyola, Sérgio Vitorino Cardoso, Leandro Alves Neves, Paulo Rogério de Faria e Marcelo Zanchetta do Nascimento (2024a). “OralEpitheliumDB: A dataset for oral epithelial dysplasia image segmentation and classification”. Em: *Journal of Imaging Informatics in Medicine*, pp. 1–20.
- (ago. de 2024b). “OralEpitheliumDB: A Dataset for Oral Epithelial Dysplasia Image Segmentation and Classification”. Em: *Journal of Imaging Informatics in Medicine*

- 37.4, pp. 1691–1710. ISSN: 2948-2933. DOI: <10.1007/s10278-024-01041-w>. URL: <https://doi.org/10.1007/s10278-024-01041-w>.
- Silva, Adriano Barbosa, Alessandro Santana Martins, Thaína Aparecida Azevedo Tosta, Leandro Alves Neves, João Paulo Silva Servato, Marcelo Sivieri de Araújo, Paulo Rogério de Faria e Marcelo Zanchetta do Nascimento (2022). “Computational analysis of histological images from hematoxylin and eosin-stained oral epithelial dysplasia tissue sections”. Em: *Expert Systems with Applications* 193, p. 116456. ISSN: 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2021.116456>. URL: <https://www.sciencedirect.com/science/article/pii/S0957417421017401>.
- Silva, Adriano Barbosa, Guilherme B. Rozendo, Thaína A. A. Tosta, Alessandro S. Martins, Adriano M. Loyola, Sérgio V. Cardoso, Alessandra Lumini, Leandro A. Neves, Paulo R. de Faria e Marcelo Z. do Nascimento (2023). “CNN Ensembles for Nuclei Segmentation on Histological Images of OED”. Em: *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 601–604. DOI: <10.1109/CBMS58004.2023.00286>.
- Silva, Adriano Barbosa, Dalí Freire Dias dos dos Santos, Thaína A. A. Tosta, Alessandro S. Martins, Leandro A. Neves, Bruno A. N. Travenço, Paulo R. de Faria e Marcelo Z. do Nascimento (2020). “Segmentation of Oral Epithelial Dysplasias Employing Mask R-CNN and Color Normalization”. Em: *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 2818–2824. DOI: <10.1109/BIBM49941.2020.9313101>.
- Silva, Adriano Barbosa, Thaína A. A. Tosta, Leandro A. Neves, Alessandro S. Martins, Paulo R. De Faria e Marcelo Z. Do Nascimento (2024c). “Ensemble of Semantic Segmentation Models for Oral Epithelial Dysplasia Images”. Em: *2024 37th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp. 1–6. DOI: <10.1109/SIBGRAPI62404.2024.10716304>.
- Silva, André Fernando Quaresma da (2024). “Normalização de Cores de Imagens Histológicas Coradas com H&E usando Aprendizagem de Dicionário”. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação). São José dos Campos: Universidade Federal de São Paulo. URL: <https://repositorio.unifesp.br/items/141bf99b-ceb7-4cb7-b42b-ce9fc41c29b5>.
- Smith, Joel, Tim Rattay, Chris McConkey, Tim Helliwell e Hisham Mehanna (2009). “Biomarkers in dysplasia of the oral cavity: a systematic review”. Em: *Oral oncology* 45.8, pp. 647–653.
- Solomon, Chris e Toby Breckon (2000). *Fundamentos de processamento digital de imagens: uma abordagem prática com exemplos em Matlab*. Grupo Gen-LTC.
- Speight, Paul M., Timothy J. Abram, Pierre N. Floriano, Robert James, Julie Vick, Martin H. Thornhill, Craig Murdoch, Christine Freeman, Anne M. Hegarty, Katy D’Apice, A. Ross Kerr, Joan Phelan, Patricia Corby, Ismael Khouly, Nadarajah Vigneswaran,

- Jerry Bouquot, Nagi M. Demian, Y. Etan Weinstock, Spencer W. Redding, Stephanie Rowan, Chih-Ko Yeh, H. Stan McGuff, Frank R. Miller e John T. McDevitt (2015). “Interobserver agreement in dysplasia grading: toward an enhanced gold standard for clinical pathology trials”. Em: *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology* 120.4, 474–482.e2. ISSN: 2212-4403.
- Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever e Ruslan Salakhutdinov (2014). “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”. Em: *Journal of Machine Learning Research* 15.56, pp. 1929–1958.
- Tilakaratne, Wanninayake M, Primali R Jayasooriya, Nadeena S Jayasuriya e Rohana Kumara De Silva (2019). “Oral epithelial dysplasia: Causes, quantification, prognosis, and management challenges”. Em: *Periodontology 2000* 80.1, pp. 126–147.
- Tompson, Jonathan, Ross Goroshin, Arjun Jain, Yann LeCun e Christopher Bregler (2015). *Efficient Object Localization Using Convolutional Networks*. eprint: <1411.4280>.
- Tosic, Ivana e Pascal Frossard (2011). “Dictionary Learning”. Em: *IEEE signal processing magazine* 28, pp. 27–38.
- Tosta, Thaína A. Azevedo, Paulo Rogério de Faria, Leandro Alves Neves, Alessandro Santana Martins, Chetna Kaushal e Marcelo Zanchetta do Nascimento (fev. de 2024). “Evaluation of sparsity metrics and evolutionary algorithms applied for normalization of H&E histological images”. Em: *Pattern Analysis and Applications* 27.1, p. 11. ISSN: 1433-755X. DOI: <10.1007/s10044-024-01218-7>. URL: <<https://doi.org/10.1007/s10044-024-01218-7>>.
- TOSTA, Thaína Aparecida Azevedo, Paulo Rogério de Faria, Leandro Alves Neves e Marcelo Zanchetta do Nascimento (2019a). “Color normalization of faded HE-stained histological images using spectral matching”. Em: *Computers in Biology and Medicine* 111.
- (2019b). “Computational normalization of HE-stained histological images: Progress, challenges and future potential”. Em: *Artificial Intelligence in Medicine* 95, pp. 118–132.
- TOSTA, Thaína Aparecida Azevedo, Paulo Rogério de Faria, João Paulo Silva Servato, Leandro Alves Neves, Guilherme Freire Roberto, Alessandro Santana Martins e Marcelo Zanchetta do Nascimento (2019c). “Unsupervised method for normalization of hematoxylin-eosin stain in histological images”. Em: *Computerized Medical Imaging and Graphics* 77.
- Vahadane, Abhishek, Tingying Peng, Amit Sethi, Shadi Albarqouni, Lichao Wang, Maximilian Baust, Katja Steiger, Anna Melissa Schlitter, Irene Esposito e Nassir Navab (2016). “Structure-Preserving Color Normalization and Sparse Stain Separation for Histological Images”. Em: *IEEE Transactions on Medical Imaging* 35.

- Warnakulasuriya, Saman, Jesper Reibel, J Bouquot e Erik Dabelsteen (2008). “Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement”. Em: *Journal of Oral Pathology & Medicine* 37.3, pp. 127–133.
- Wei, Lisheng, Quan Gan e Tao Ji (2017). “Cervical cancer histology image identification method based on texture and lesion area features”. Em: *Computer Assisted Surgery* 22.sup1, S258–S267. DOI: <10.1080/24699322.2017.1389397>.
- Wild, C. P., E. Weiderpass e B. W. Stewart (2020). “World cancer report: cancer research for cancer prevention”. Em: *International Agency for Research on Cancer* 1.
- Xie, Xinpeng, Yuexiang Li, Menglu Zhang e Linlin Shen (2019). “Robust Segmentation of Nucleus in Histopathology Images via Mask R-CNN”. Em: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Cham: Springer International Publishing, pp. 428–436. ISBN: 978-3-030-11723-8.
- Yanase, Juri e Evangelos Triantaphyllou (2019). “A systematic survey of computer-aided diagnosis in medicine: Past and present developments”. Em: *Expert Systems with Applications* 138, p. 112821.
- Yang, Dong, Daguang Xu, S Kevin Zhou, Bogdan Georgescu, Mingqing Chen, Sasa Grbic, Dimitris Metaxas e Dorin Comaniciu (2017). “Automatic liver segmentation using an adversarial image-to-image network”. Em: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 507–515.
- Yeasmin, Most Nilufa, Md Al Amin, Tasmim Jamal Joti, Zeyar Aung e Mohammad Abdul Azim (2024). “Advances of AI in image-based computer-aided diagnosis: A review”. Em: *Array* 23, p. 100357. ISSN: 2590-0056. DOI: <<https://doi.org/10.1016/j.array.2024.100357>>.
- Zelst, Jan CM van, Tao Tan, Ritse M Mann e Nico Karssemeijer (2020). “Validation of radiologists findings by computer-aided detection (CAD) software in breast cancer detection with automated 3D breast ultrasound: a concept study in implementation of artificial intelligence software”. Em: *Acta Radiologica* 61.3, pp. 312–320.
- Zhang, Lin, Lei Zhang, Xuanqin Mou e David Zhang (2011). “FSIM: A Feature Similarity Index for Image Quality Assessment”. Em: *IEEE Transactions on Image Processing* 20.8, pp. 2378–2386.
- Zhao, Zhi-Qi, Peng Zheng, Shou-Tao Xu e Xindong Wu (2019). “Object detection with deep learning: A review”. Em: *IEEE transactions on neural networks and learning systems* 30.11, pp. 3212–3232.