



UNIVERSIDADE FEDERAL DO MARANHÃO
Programa de Pós-Graduação em Ciência da Computação

Marcos Gabriel Mendes Lauande

***Classificação de Imagens Histopatológicas de Câncer de Pênis
usando Redes Convolucionais e Transferência de Aprendizagem***

São Luís
2022

Marcos Gabriel Mendes Lauande

Classificação de Imagens Histopatológicas de Câncer de Pênis usando Redes Convolucionais e Transferência de Aprendizagem

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Ciência da Computação, ao Programa de Pós-Graduação em Ciência da Computação, da Universidade Federal do Maranhão.

Programa de Pós-Graduação em Ciência da Computação

Universidade Federal do Maranhão

Orientador: Prof. Dr. Geraldo Braz Júnior

Coorientador: Prof. Dr. João Dallyson Sousa de Almeida

São Luís - MA

2022

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Diretoria Integrada de Bibliotecas/UFMA

Lauande, Marcos Gabriel Mendes.

Classificação de Imagens Histopatológicas de Câncer de Pênis usando Redes Convolucionais e Transferência de Aprendizagem / Marcos Gabriel Mendes Lauande. - 2022.
67 f.

Coorientador(a): João Dallyson Sousa de Almeida.

Orientador(a): Geraldo Braz Júnior.

Dissertação (Mestrado) - Programa de Pós-graduação em Ciência da Computação/ccet, Universidade Federal do Maranhão, São Luís, MA, 2022.

1. Câncer de pênis. 2. Densenet. 3. Imagens histopatológicas. 4. Redes neurais convolucionais. 5. Transferência de aprendizado. I. de Almeida, João Dallyson Sousa. II. Júnior, Geraldo Braz. III. Título.

Marcos Gabriel Mendes Lauande

Classificação de Imagens Histopatológicas de Câncer de Pênis usando Redes Convolucionais e Transferência de Aprendizagem

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Ciência da Computação, ao Programa de Pós-Graduação em Ciência da Computação, da Universidade Federal do Maranhão.

Trabalho aprovado. São Luís - MA, 24 de Junho de 2022:

Prof. Dr. Geraldo Braz Júnior
Orientador

Prof. Dr. João Dallyson Sousa de Almeida
Coorientador

Prof. Dr. Daniel Lima Gomes Junior
Examinador Interno

Prof. Dr. Rui Miguel Gil da Costa Oliveira
Examinador Externo

Prof. Dra. Fátima de Lourdes dos Santos Nunes Marques
Examinador Externo

São Luís - MA
2022

Aos meus imprescindíveis pais por sempre me incentivarem a estudar.

Agradecimentos

Agradeço a Deus pela força e fé para atingir meus objetivos.

Agradeço a minha família pelo apoio e motivação, não só no período do mestrado, mas como em todos os momentos da minha vida.

Agradeço ao meu orientador Geraldo Braz Júnior, ao meu co-orientador João Dallyson de Sousa Almeida, aos docentes do Programa de Pós-graduação em Ciência da Computação e Programa de Pós-graduação em Saúde do Adulto pelos direcionamentos, paciência, disponibilidade e motivação por toda essa minha trajetória acadêmica no mestrado que resultou neste trabalho.

Agradeço também aos meus colegas do programa e dentre outros pesquisadores pelos conhecimentos compartilhados.

*"Oh! vós, que tendes tempo sem ter conta, Não gasteis vosso tempo em passa-tempo.
Cuidai, enquanto é tempo em vossa conta." (Frei Antonio das Chagas)*

Resumo

O câncer de pênis é um tumor raro que corresponde a 2% dos casos de câncer nos homens no Brasil, sendo que o Maranhão é um dos estados com mais incidência. Análises histopatológicas são comumente usadas no seu diagnóstico, possibilitando a avaliação do grau da enfermidade, a sua evolução e natureza. Há aproximadamente uma década, trabalhos científicos na área de Aprendizagem Profunda têm sido desenvolvidos para auxiliar médicos patologistas a tomarem decisões de maneira rápida e confiável, abrindo possibilidades de novas contribuições para aperfeiçoamento de uma atividade tão complexa e demorada para estes profissionais. Neste trabalho, o desenvolvimento de um método baseado em aprendizado profundo e transferência de aprendizado é apresentado. Estes conceitos foram aplicados utilizando uma rede convolucional do tipo DenseNet para o diagnóstico do câncer de pênis em um *dataset* de 194 imagens histopatológicas construído pelo Projeto Câncer de Pênis da Amazônia Legal. Além disso, foi avaliado outro método baseado em *deep features* para fins comparativos. Utilizou-se, também, o algoritmo Grad-CAM para explicar o modelo gerado com o objetivo de descobrir quais as regiões das imagens foram mais relevantes durante o processo de classificação. Nos experimentos realizados, para diferenciar imagens histopatológicas normais e aquelas com carcinoma de células escamosas, foram alcançados um F1-Score de até 97,39% e uma sensibilidade de até 98,33%.

Palavras-chave: Câncer de pênis, imagens histopatológicas, redes neurais convolucionais, Densenet, transferência de aprendizado.

Abstract

Penile cancer is a rare tumor that accounts for 2% of cancer cases in men in Brazil, with Maranhão being one of the states with the highest incidence. Histopathological analyzes are commonly used in its diagnosis, making it possible to assess the degree of the disease, its evolution, and its nature. For approximately a decade, scientific works in Deep Learning have been developed to help pathologists make decisions quickly and reliably, opening up possibilities for new contributions to improve such a complex and time-consuming activity for these professionals. In this work, the development of a method based on deep learning and transfer of learning is presented. These concepts were applied using a DenseNet-type convolutional network to diagnose penile cancer in a dataset of histopathological with 194 images built by the Legal Amazon Penis Cancer Project. In addition, another method based on deep features was evaluated for comparative purposes. The Grad-CAM algorithm was also used to explain the generated model to discover which regions of the images were relevant during the classification process. In the experiments performed to differentiate normal histopathological images and those with squamous cell carcinoma, an F1-Score of up to 97.39% and a sensitivity of up to 98.33% was achieved.

Keywords: Penile cancer, histopathological images, convolutional neural network, Densenet, transfer learning

Lista de ilustrações

Figura 1 – Figura da anatomia peniana externa e interna.	22
Figura 2 – Exemplo da ação da coloração com hematoxilina e eosina nos componentes do tecido peniano de lâmina histopatológica.	25
Figura 3 – No primeiro exemplo, as áreas de crescimento infiltrativo (invasivo) da lesão estão indicadas conforme as setas que descrevem sua direção. No segundo exemplo, são observadas as pérolas de queratina indicadas pelos retângulos.	26
Figura 4 – Exemplo de uma rede neural multicamadas também conhecida como <i>multi-layer perceptron</i>	29
Figura 5 – Exemplo da influência da taxa de aprendizado na convergência da rede.	30
Figura 6 – Rede convolucional clássica - LeNet5.	32
Figura 7 – Exemplo da técnica de convolução: neste caso, para preservar o tamanho da matriz, utilizou-se a adição de bordas (<i>padding</i>) com os valores iguais a zero.	33
Figura 8 – Exemplo da técnica de <i>pooling</i> , neste caso, o <i>max pooling</i> (valor máximo).	33
Figura 9 – Exemplo de uma DenseNet com três blocos densos.	34
Figura 10 – Exemplo de uma <i>dense block</i> e as conexões entre as camadas.	35
Figura 11 – Tabela proveniente do trabalho original de Huang et al. (2017) onde podem ser vistas as arquiteturas e suas configurações de camadas.	35
Figura 12 – Nesta figura é exemplificada a ocorrência do <i>overfitting</i> e <i>underfitting</i>	36
Figura 13 – Exemplos de transformações aplicadas em imagens histopatológicas no processo de <i>data augmentation</i>	38
Figura 14 – Diagrama do método proposto.	40
Figura 15 – Exemplos de imagens histopatológicas de câncer de pênis por categoria e ampliação.	41
Figura 16 – Aplicação do algoritmo de realce CLAHE em uma imagem histopatológica de câncer de pênis.	43
Figura 17 – Arquitetura DenseNet-201 adaptada ao problema de classificação binária.	44
Figura 18 – Exemplo de particionamento utilizando validação cruzada com cinco <i>folds</i>	45
Figura 19 – Diagrama do método do estudo de ablação que utiliza as técnica de <i>Deep Features</i> em uma rede convolucional pré-treinada na ImageNet. As camadas convolucionais funcionam como um extrator de características.	49
Figura 20 – Resultado do algoritmo Grad-CAM ao sobrepor os mapas de calor com as imagens de ampliação 40X avaliadas.	53

Figura 21 – Resultado do algoritmo Grad-CAM ao sobrepor os mapas de calor com as imagens de ampliação 100X avaliadas.	54
---	----

Lista de tabelas

Tabela 1	– Lista dos trabalhos relacionados obtidos através da revisão da literatura.	21
Tabela 2	– Distribuição das imagens conforme ampliação e classificação patológica.	40
Tabela 3	– Matriz de confusão para tarefas de classificação binária.	44
Tabela 4	– Comparativo de métricas entre os diversos experimentos que comprovam o desempenho do método baseado na DenseNet.	48
Tabela 5	– Comparativo de métricas em percentuais entre os diversos experimentos utilizando a arquitetura de CNN Xception.	48
Tabela 6	– Comparativo de métricas em percentuais entre os diversos experimentos utilizando a arquitetura de CNN Inception ResNet V2.	49
Tabela 7	– Tabela que descreve os parâmetros disponíveis no espaço de busca do algoritmo grid-search para cada classificador utilizado nos experimentos.	50
Tabela 8	– Resultados da técnica <i>Deep Features</i> com a DenseNet.	50
Tabela 9	– Resultados da técnica <i>Deep Features</i> com a Inception ResNet.	51
Tabela 10	– Resultados da técnica <i>Deep Features</i> com a Xception.	51
Tabela 11	– Melhores parâmetros encontrados pelo algoritmo grid-search para os melhores resultados dos experimentos realizados.	52
Tabela 12	– Comparação entre o método proposto (linha 1 e 2) e o método baseado em <i>Deep Features</i> desenvolvido pelo mesmo autor (linha 3 e 4).	52
Tabela 13	– Artigo publicado em evento científico em relação ao tema diagnóstico de câncer de pênis utilizando imagens histopatológicas.	56

Lista de abreviaturas e siglas

ACS	<i>American Cancer Society</i>
AHE	<i>Adaptative Histogram Equalization</i>
CLAHE	<i>Contrast Limited Adaptive Histogram Equalization</i>
CLBP	<i>Completed Local Binary Pattern Gray-Level</i>
GLCM	<i>Gray-Level Co-Occurrence Matrices</i>
CNN	<i>Convolutional Neural Network</i>
DT	<i>Decision Tree</i>
FC	<i>Fully Connected</i>
GANs	<i>Generative Adversarial Networks</i>
GCN	<i>Global Contrast Normalization</i>
Grad-CAM	<i>Gradient-weighted Class Activation Mapping</i>
HE	<i>Histogram Equalization</i>
HPV	<i>Human Papiloma Virus</i>
IA	<i>Inteligência Artificial</i>
IBGE	<i>Instituto Brasileiro de Geografia e Estatística</i>
INCA	<i>Instituto Nacional do Câncer</i>
KNN	<i>K-Nearest Neighborhood</i>
LBP	<i>Local Binary Patterns</i>
LPQ	<i>Local Phase Quantization</i>
NCI	<i>National Cancer Institute</i>
PET	<i>Positron Emission Tomography</i>
PFTAS	<i>Parameter-Free Threshold Adjacency Statistics</i>
PWPT	<i>Pyramid-Structure Wavelet Transform</i>

ORB	<i>Oriented FAST and Rotated BRIEF</i>
RF	<i>Random Forest</i>
SVM	<i>Support Vector Machine</i>
TWT	<i>Tree-Structure Wavelet Transform</i>
VGG	<i>Visual Geometry Group</i>
WSI	<i>Whole Slide Image</i>

Sumário

1	INTRODUÇÃO	15
1.1	Objetivos	16
1.1.1	Objetivos Específicos	16
1.1.2	Contribuições	17
1.2	Organização do Trabalho	17
2	TRABALHOS RELACIONADOS	18
3	FUNDAMENTAÇÃO TEÓRICA	22
3.1	Câncer de Pênis	22
3.2	Análise Histopatológica	23
3.3	Histopatologia Digital	26
3.4	Análise de Imagens Médicas Auxiliadas por Computador	27
3.5	Redes Neurais e Aprendizado Profundo	28
3.5.1	Redes Neurais Convolucionais	31
3.5.2	O Problema do Vanishing Gradient nas Redes Neurais Profundas	33
3.5.3	Arquitetura DenseNet	34
3.5.4	Overfitting e Underfitting	36
3.6	Transferência de Aprendizado e Fine-tuning	37
3.7	Data Augmentation	37
3.8	Considerações Finais	39
4	METODOLOGIA	40
4.1	Aquisição das imagens histopatológicas	40
4.2	Pré-processamento	41
4.3	Construção e Treinamento da Rede Neural Convolucional	42
4.4	Avaliação	44
4.5	Considerações Finais	46
5	RESULTADOS	47
5.1	Resultados dos Experimentos com o Método Proposto	47
5.2	Estudo de Ablação com a Técnica Deep Features	49
5.3	Explicabilidade do Modelo Gerado pelo Método Proposto	52
6	CONCLUSÃO	55

REFERÊNCIAS 57

1 Introdução

O câncer está relacionado ao crescimento desordenado de células. Dependendo do grau de malignidade, pode invadir tecidos adjacentes ou órgãos levando o paciente à morte se não houver um tratamento precoce adequado. O câncer de pênis é um tumor raro e representa 0,4% a 0,6% de todos os cânceres na Europa e na América do Norte, mas é consideravelmente mais comum em países em desenvolvimento da América Latina, África e Ásia (DOUGLAWI; MASTERSON, 2017). Conforme o INCA (2021), esta doença tem maior incidência em homens a partir de 50 anos no Brasil, sendo mais comum nas regiões Norte e Nordeste do país e corresponde a 2% da incidência de câncer nos homens. Um relatório proveniente de dados coletados em um dos estados do nordeste brasileiro com maior incidência desta doença no mundo, o Maranhão, indica taxa padronizada por idade de 6,15 por 100 mil (COELHO et al., 2018).

Com base nos indicadores do IBGE (2022), o Maranhão é um dos estados mais pobres do Brasil, com renda per capita de R\$ 635,00 e Índice de Desenvolvimento Humano de 0,639. Somados a esses fatores, a alta taxa de habitantes rurais, a distância dos grandes centros de saúde e o baixo nível de escolaridade fornecem as características que propiciam o desenvolvimento do câncer peniano (COELHO et al., 2018). Além disso, há uma alta prevalência do papilomavírus humano (*Human Papiloma Virus*, ou HPV) na população geral do estado e relatos de casos da prática de zoofilia entre os pacientes dessa região. Um estudo de Vieira et al. (2020), baseado em um grupo de 116 pacientes tratados em hospitais da capital, São Luís, informa que a proporção de realizações de penectomia (amputação peniana) neste estado supera a de pacientes submetidos a esta cirurgia em outras regiões do país. Estas amputações são estreitamente relacionadas à procura tardia de tratamento especializado, consequência dos fatores citados acima. Isso é muito preocupante, pois tratamentos tão invasivos podem não só deixar sequelas físicas, mas também impactos psicológicos, sexuais e sociais.

Conforme (ACS, 2021) e (THOMAS et al., 2021), um dos exames que pode ser indicado para o diagnóstico dessa doença é a análise histopatológica de tecidos coletados através de biópsia, que consiste na avaliação microscópica de tecidos bem finos extraídos da região de interesse. No entanto, de acordo com Melo et al. (2020), esta atividade tende a ser muito complexa e demorada, por causa do processo ser claramente minucioso e de conclusão subjetiva, dependendo muito da expertise do patologista. À vista disso, soluções baseadas em software podem auxiliar estes profissionais trazendo mais confiabilidade, possibilitando um diagnóstico precoce, sendo que quanto mais cedo a doença for detectada e juntamente com um tratamento adequado, maior a probabilidade de cura.

Com a evolução de técnicas de Aprendizado de Máquina e processamento de imagens aplicadas a imagens médicas, várias pesquisas científicas correlatas baseadas em diagnóstico utilizando imagens histopatológicas têm sido desenvolvidas (SRINIDHI; CIGA; MARTEL, 2021), por exemplo, para o câncer de mama (CRUZ-ROA et al., 2014) (VIDYARTHI et al., 2019), próstata (LINKON et al., 2021), fígado (KIANI et al., 2020), cólon (SARWINDA et al., 2021) e pulmão (WEI et al., 2019). Dessa forma, as arquiteturas de Aprendizado Profundo, especificamente as redes neurais convolucionais, acabaram tornando-se bastante populares para este tipo de problema. Essa popularidade deu-se devido ao seu desempenho, em determinados casos, superior aos métodos tradicionais de extração de características e aos classificadores não baseados em redes neurais. Isso pode ser notado nos resultados obtidos em Malavade, Melinamath e Pardeshi (2018), Srinidhi, Ciga e Martel (2021) e Zhou et al. (2020), por exemplo. Contudo, em relação ao diagnóstico de câncer baseado em imagens a níveis celulares do tecido peniano, a falta de um *dataset* que possa subsidiar os experimentos dos pesquisadores se torna um fator impeditivo para que métodos especializados possam ser desenvolvidos.

À vista disso, através da cessão de uso de um *dataset* de imagens histopatológicas por um grupo de pesquisa da Universidade Federal do Maranhão, o Projeto Câncer de Pênis da Amazônia Legal, este trabalho apresenta o desenvolvimento de um método automatizado baseado em imagens e técnicas de Aprendizado Profundo com a finalidade de diagnosticar o câncer de pênis, auxiliando o médico a tomar decisões de forma mais rápida. O autor deste trabalho espera que, a partir desta pesquisa, outras possam ser desenvolvidas com intuito de promover a saúde masculina utilizando tecnologias auxiliadas por computador.

1.1 Objetivos

O objetivo geral deste trabalho consiste em desenvolver um método baseado em redes neurais convolucionais e transferência de aprendizado para diagnosticar câncer de pênis em imagens histopatológicas com a finalidade de auxiliar os patologistas nesta decisão.

Além disso, o método desenvolvido foi validado em um *dataset* construído por um grupo de pesquisadores do Projeto Câncer de Pênis da Amazônia Legal, sendo composto por 194 imagens histopatológicas classificadas como normal ou carcinoma de células escamosas (*squamous cell carcinoma*).

1.1.1 Objetivos Específicos

Especificamente, este trabalho busca os seguintes objetivos aplicados ao problema de classificação de imagens histopatológicas de câncer de pênis:

- Avaliar abordagens de Aprendizado Profundo utilizando as técnicas de transferência de aprendizado, *fine-tuning* e *deep features* para o diagnóstico de câncer de pênis;
- Realizar um comparativo do desempenho dos métodos desenvolvidos na base de imagens do Projeto Câncer de Pênis da Amazônia Legal;
- Avaliar a explicabilidade dos modelos utilizando o algoritmo *Gradient-weighted Class Activation Mapping* (Grad-CAM) no intuito de verificar quais regiões das imagens foram mais importantes no processo de predição;

1.1.2 Contribuições

Destacam-se como principais contribuições:

- A proposição de um método para o diagnóstico automatizado de câncer de pênis que utiliza transferência de aprendizado em uma rede convolucional do tipo DenseNet em imagens de microscópio de diferentes níveis de ampliação;
- A construção de um *dataset* de imagens histopatológicas de câncer de pênis para o diagnóstico automatizado auxiliado por computador;

1.2 Organização do Trabalho

Este trabalho está estruturado da seguinte forma:

- O Capítulo 2 descreve trabalhos relacionados à análise de imagens histopatológicas que utilizam diversas soluções baseadas em Aprendizado de Máquina e processamento de imagens.
- O Capítulo 3 trata da fundamentação teórica no qual são abordados conceitos referentes ao câncer de pênis, análise histopatológica, patologia digital, tecnologias baseadas em Aprendizado Profundo e as técnicas utilizadas para melhorar seu desempenho no processo de treinamento e classificação.
- O Capítulo 4 apresenta o método desenvolvido e suas etapas para o diagnóstico de câncer de pênis baseado em Aprendizado Profundo.
- O Capítulo 5 trata sobre os resultados obtidos e discussões em relação aos experimentos realizados utilizando o método proposto, além do estudo de ablação com a técnica *deep features* e a análise via algoritmo Grad-CAM para fins de explicabilidade do modelo gerado pela rede convolucional adotada.
- O Capítulo 6 apresenta as considerações finais sobre os resultados e trabalhos futuros e os artigos científicos resultantes desta pesquisa.

2 Trabalhos Relacionados

O reconhecimento por computador de imagens médicas, especificamente as referentes às histopatológicas de câncer, é um tópico de pesquisa bastante explorado, tendo muitos trabalhos disponíveis para obtenção de fundamentação teórica que serviram de base para o desenvolvimento do método deste trabalho.

Os trabalhos abordados utilizam tanto metodologias clássicas de processamento de imagens, que remetem à extração de características e depois o uso de um algoritmo classificador, quanto abordagens que se valem de Aprendizagem Profunda.

Filipczuk et al. (2013) utilizaram a Transformada de Hough para detectar os núcleos de células em imagens histopatológicas que tiveram suas características extraídas e utilizadas como entrada para um classificador *Support Vector Machine* (SVM). Spanhol et al. (2016b) comparam técnicas de extração de características como *Local Binary Patterns* (LBP), *Completed Local Binary Pattern* (CLBP), *Gray-Level Co-Occurrence Matrices* (GLCM), *Parameter-Free Threshold Adjacency Statistics* (PFTAS), *Oriented FAST and Rotated BRIEF* (ORB) e *Local Phase Quantization* (LPQ) em experimentos com classificadores não baseados em redes neurais pela primeira vez em um *dataset* composto por 7.909 imagens de câncer de mama chamado BreakHis. Samah, Fauzi e Mansor (2017) propuseram um classificador de imagens histopatológicas de câncer de mama que utiliza um classificador *K-Nearest Neighbors* (KNN) e os descritores de imagens GLCM, LBP, *Tree-Structure Wavelet Transform* (TWT) e *Pyramid-Structure Wavelet Transform* (PWPT).

Com o passar do tempo, técnicas tradicionais de extração de características foram sendo substituídas por técnicas baseadas em Aprendizado Profundo para este tipo de problema, apresentando melhores resultados. Em Sharma et al. (2017), uma arquitetura de CNN foi construída e aplicada para a classificação e detecção de necrose utilizando um *dataset* de imagens histopatológicas de câncer gástrico (BEHRENS et al., 2015). Esse trabalho mostrou um desempenho superior em comparação aos métodos tradicionais que neste caso utilizaram o classificador *Random Forest*. Já em Spanhol et al. (2016a), os autores apresentaram um método que utiliza uma rede convolucional baseada na AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012). O método foi comparado com os resultados do seu trabalho anterior (SPANHOL et al., 2016b), que foi baseado em técnicas de extração de características e classificação não baseados em redes neurais, mostrando-se superior.

Ao realizar essa revisão de literatura, percebeu-se que o uso de técnicas de pré-processamento é muito importante para melhorar o desempenho de tarefa de classificação

de imagens de natureza histopatológica. Sarwinda et al. (2021) demonstraram a utilização do algoritmo *Contrast Limited Histogram Equalization* (CLAHE) juntamente com as redes neurais ResNet-18 e ResNet-50 para um problema de classificação binária (maligno ou benigno). Neste caso, uma base de imagens de tecidos do intestino grosso teve suas amostras transformadas para a escala de cinza e pré-processadas através do algoritmo CLAHE, melhorando substancialmente o contraste das imagens e os resultados da classificação. Em Vidyarthi et al. (2019), este mesmo algoritmo foi utilizado no melhoramento de imagens histopatológicas de câncer de mama. Após este pré-processamento, as imagens foram transformadas pelo algoritmo *Watershed* (DIGABEL; LANTUÉJOU, 1978), neste caso, a ideia-chave por trás do algoritmo é transformar a imagem em uma outra cujas bacias de captação são os objetos que precisam ser identificados. A partir deste processo, as imagens são utilizadas como entrada para uma rede convolucional LeNet-5 (LECUN et al., 1989). Em comparação com o experimento sem a etapa de pré-processamento (CLAHE+ *Watershed*), o método apresentado conseguiu resultados bastante expressivos.

Além de problemas envolvendo iluminação e contraste, conforme Roy et al. (2018), dependendo da forma como foi feita a aquisição das imagens ou do processo de coloração, as imagens tendem a ter o conjunto de cores diferentes entre amostras, afetando o desempenho de métodos baseados em Aprendizado de Máquina. Ciompi et al. (2017) demonstraram uma normalização na coloração apresentada nas imagens histopatológicas de câncer colorretal. Com base em uma amostra modelo, as outras amostras foram transformadas com base na distribuição de cores deste padrão. As imagens transformadas serviram de entrada para uma rede convolucional baseada no trabalho de Simonyan e Zisserman (2014). Os resultados dos experimentos com o método com a normalização das cores foram muito significativos em relação aos experimentos sem essa transformação nas imagens. Em Anghel et al. (2019) foi apresentado um método de normalização. Desta vez, o processo de transformação das imagens foi baseado no uso de um método não supervisionado que estima a paleta de cores das imagens histopatológicas. Este mecanismo otimizado integra uma técnica de limiarização para determinar os *pixels* úteis e usa um método de amostragem de *pixels* que reduz significativamente o tempo de processamento do algoritmo de normalização. As imagens normalizadas serviram de entrada para uma rede VGG (*Visual Geometry Group*) adaptada, produzindo resultados notáveis.

No contexto de técnicas de Aprendizagem Profunda, o uso de transferência de aprendizado na classificação de imagens histopatológicas se tornou muito promissora e popular; principalmente quando os experimentos estão restritos a computadores menos poderosos e os *datasets* com poucas amostras. Este procedimento consiste na utilização de redes pré-treinadas em determinado domínio em um outro semelhante ou diferente. A aplicação desta técnica pode ser observada no artigo de Spanhol et al. (2017), no qual os autores utilizaram as camadas convolucionais de uma rede CaffeNet pré-treinada no *dataset* ImageNet (RUSSAKOVSKY et al., 2015) como um extrator de características em que

sua saída serviu de entrada para um classificador de Regressão Logística. Boumaraf et al. (2021) construíram um método de diagnóstico de câncer de mama baseado em transferência de aprendizado que utiliza o algoritmo *Global Contrast Normalization* (GCN) para pré-processar as imagens que serviram de entrada para uma rede ResNet-18 pré-treinada no *dataset* ImageNet. Choudhary et al. (2021) propuseram um método de diagnóstico de câncer de mama adaptado para dispositivos com pouco poder de processamento, no qual uma rede pré-treinada teve seus pesos menos importantes suprimidos de forma a aprimorar seu desempenho em experimentos no *dataset* IDC (BOLHASANI et al., 2020).

Além dos métodos de transferência de aprendizado, o *data augmentation* pode ser usado para contornar a limitação da quantidade de amostras, o desbalanceamento de classes e o problema de *overfitting*, que podem afetar consideravelmente o desempenho das técnicas de Aprendizado Profundo, já que estas performam melhor em bases de imagens com milhares ou milhões de amostras, como a ImageNet. Essa técnica faz com que as imagens sofram algum tipo de transformação, produzindo, assim, novas amostras. Tellez et al. (2019) apresentam um método baseado em redes convolucionais no qual as imagens histopatológicas (tecidos provenientes de quatro órgãos) passam pelo processo de *augmentation* através de transformações morfológicas, baseadas em mudanças de espaços de cor, brilho e contraste apresentando resultados promissores. Rakhlin et al. (2018) aplicaram a técnica *color augmentation*, que consiste em fazer perturbações aleatórias no espaço de cor RGB (*Red, Green and Blue*) das imagens histopatológicas de câncer de mama do *dataset* BACH (ARAUJO et al., 2017). Estas amostras serviram de entrada para uma rede VGG-16 que extraiu as características para serem utilizadas no classificador LightGBM. Faryna, Laak e Litjens (2021) propuseram método baseado em redes convolucionais que aplica uma solução automatizada baseada no algoritmo RandAugment (CUBUK et al., 2020) com o objetivo de descobrir boas práticas de *data augmentation* para imagens histopatológicas, resultando em um desempenho superior ao trabalho de Tellez et al. (2019).

Os trabalhos abordados nesta seção estão resumidos na Tabela 1. Estes representam uma amostra da diversidade crescente de estudos baseados em Aprendizado de Máquina com a finalidade de automatizar a atividade de análise de imagens histopatológicas. Neste trabalho, desenvolvemos um método baseado nos conceitos de Aprendizado Profundo e na literatura revisada, via especialização da rede neural convolucional, fundamentado em métodos de pré-processamento e transferência de aprendizado para o diagnóstico de câncer em imagens histopatológicas de câncer de pênis. A fundamentação teórica usada para o desenvolvimento deste método, assim como o método em si, é assunto das seções seguintes deste trabalho.

Tabela 1 – Lista dos trabalhos relacionados obtidos através da revisão da literatura.

Autor/Dataset	Método	Resultado
Filipcuk et al. (2013) Imagens cedidas pelo Hospital Regional em Zielona Góra, Polônia.	Otsu + Transformada de Hough + SVM	Acurácia de até 98,43%
Spanhol et al. (2016b) BreakHis	PFTAS + SVM	Taxa de Reconhecimento de até 85,1%
Samah, Fauzi e Mansor (2017) BreakHis	Descritores diversos + K-Nearest Neighbors (KNN)	Acurácia geral de até 85,62%
Sharma et al. (2017) dataset de Vidyarthi et al. (2019)	Augmentation + AlexNet modificada	Acurácia de até 81,44%
Spanhol et al. (2016a) BreakHis	Augmentation + CaffeNet modificada	Acurácia de até 89,6%
Sarwinda et al. (2021) Warwick-QU	CLAHE + Resnet-18/50	Acurácia de até 88,0%
Vidyarthi et al. (2019) BreakHis	CLAHE/Watershed + LeNet5	Acurácia de até 90,39%
Ciampi et al. (2017) dataset de Kather et al. (2016)	Normalização das cores + ConvNet	Acurácia de até 75,55%
Anghel et al. (2019) Camelyon16	Normalização das cores + VGG adaptada	F1-Score de até 87%
Spanhol et al. (2017) BreakHis	CaffeNet + regressão logística	Acurácia de até 86,3%
Boumaraf et al. (2021) BreakHis	Global Contrast Normalization + ResNet-18	F1-Score de até 99,44%
Choudhary et al. (2021) dataset de Bolhasani et al. (2020)	ResNet-50 + filter pruning	Acurácia de até 92,07%
Tellez et al. (2019) Múltiplas bases	Augmentation + ConvNet	AUC de até 0,996
Rakhlin et al. (2018) BACH	Augmentation + model fusion (ResNet-50, InceptionV3 e VGG-16)	Acurácia de até 87,2%
Faryna, Laak e Litjens (2021) Camelyon17	RandAugment + ConvNet	AUC de até 0,964

Fonte: elaborado pelo autor

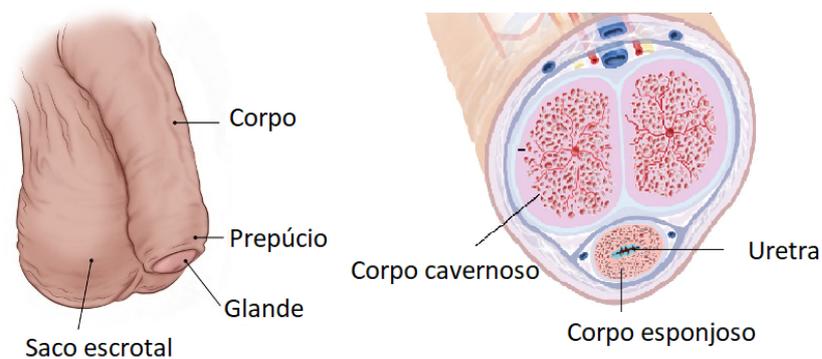
3 Fundamentação Teórica

Este capítulo apresenta os conceitos explorados para o desenvolvimento da pesquisa que resultou em um método automatizado de diagnóstico de câncer de pênis baseado em Aprendizado Profundo.

3.1 Câncer de Pênis

O pênis é o órgão sexual externo masculino e também faz parte do sistema urinário. É composto por muitos tipos de tecidos corporais, incluindo pele, nervos, músculo liso e vasos sanguíneos (ACS, 2021). Como pode ser visto na Figura 1, a região da extremidade do corpo do pênis é chamada de glande. Ao nascer, a glande é coberta por um pedaço de pele chamado prepúcio. O prepúcio é frequentemente removido em um procedimento cirúrgico chamado de circuncisão. Além da anatomia externa, o pênis possui três câmaras internas, ambas de forma cilíndrica: dois corpos cavernosos e um corpo esponjoso. Este último expandido forma a glande e cobre o canal da uretra.

Figura 1 – Figura da anatomia peniana externa e interna.



Fonte: adaptado (ACS, 2021) (NCI, 2022)

O câncer de pênis pode ser conceituado como uma doença causada pelo crescimento anormal e desordenado das células que fazem parte dos tecidos da região peniana. Esta doença é bastante temida pela população masculina devido ao seu tratamento rigoroso e por estar associada a possíveis mutilações físicas do órgão, o que pode trazer mudanças psicossociais e sociais (BARROS, 2009) (SANTOS et al., 2018). Este tipo de tumor está atrelado a alguns fatores, tais como a falta de higiene, tabagismo, infecção por papilomavírus humano (HPV), a presença de fimose e comportamento sexual de risco. O câncer de pênis possui os seguintes tipos: melanoma, carcinoma de células basais, adenocarcinoma, sarcoma, carcinoma escamoso (ou espinocelular) (ACS, 2021). Sendo

este último, o mais predominante, correspondendo a 95% dos casos no mundo e que neste trabalho representa todas as imagens histopatológicas categorizadas como câncer de pênis pertencentes ao *dataset* construído nesta pesquisa. De acordo com ACS (2021), o carcinoma escamoso é uma neoplasia epitelial maligna composta por células escamosas atípicas, podendo ter a presença de pontes intercelulares e formação de pérolas de queratina. Pode começar em qualquer lugar do pênis, tendendo a crescer lentamente. Se forem encontrados em um estágio inicial, geralmente podem ser curados.

O INCA (2021) informa que as manifestações clínicas mais comuns do câncer de pênis são “uma ferida ou úlcera persistente, ou também uma tumoração localizada na glândula, prepúcio ou corpo do pênis. A presença de um desses sinais, associados a uma secreção branca (esmegma), pode ser um indicativo de câncer no pênis. Nesses casos, é necessário consultar um especialista. Além da tumoração no pênis, a presença de gânglios inguinais (ínguas na virilha), pode ser sinal de progressão da doença (metástase)”.

Na maioria das vezes, o paciente com câncer de pênis procura o atendimento médico tardiamente, por falta de recursos locais ou mesmo por temer o tratamento cirúrgico (POMPEO et al., 2006) (COSTA et al., 2013).

Nos estágios avançados dessa patologia, o tratamento mais utilizado é a remoção cirúrgica da região afetada, que pode evoluir para uma amputação parcial ou total do pênis. Entretanto outras técnicas podem ser abertamente discutidas nesse contexto, como radioterapia, quimioterapia e a cirurgia a laser (INCA, 2021).

De acordo com Mosconi et al. (2005), o exame físico é o ponto central da avaliação clínica do câncer de pênis, consistindo na avaliação do tumor, o que inclui o tamanho, localização, mobilidade, grau de infiltração e envolvimento dos corpos cavernosos. O exame inguinal é considerado obrigatório, pois a sobrevida está diretamente correlacionada com a presença ou ausência de metástases inguinais. Além do exame físico, os exames de imagem ressonância magnética e ultrassonografia, por exemplo, podem ser empregados. Por fim, a biópsia e exame histopatológico são indispensáveis para o diagnóstico definitivo da doença.

Destaca-se como formas de prevenção algumas ações simples, como a boa higienização do pênis, a cirurgia de fimose dependendo do caso, a utilização de preservativos nas relações sexuais, o diagnóstico precoce da doença. Estas ações citadas anteriormente são os principais tópicos das campanhas de muitas entidades de saúde pública (WIND et al., 2018) (INCA, 2021).

3.2 Análise Histopatológica

A análise histopatológica é um procedimento clínico obrigatório para o diagnóstico definitivo do câncer de pênis. Este tipo de exame pode ser conceituado como o estudo das

células e tecidos de um ser vivo para realizar o diagnóstico de alguma doença. A palavra histopatologia é originária do grego, sendo que o termo “histo” significa tecido, “pathos” doença e “logia” estudo. De acordo com Neto (2012), a interpretação histopatológica é uma etapa crítica com objetivo de estabelecer o diagnóstico de benignidade e malignidade e pode discriminar os diversos tipos e graus de câncer, bem como determinar mecanismos e percursos moleculares específicos dos tumores. Essas informações acabam sendo muito importantes para estimar o prognóstico e a proposição de um melhor tratamento.

Para a realização do estudo das estruturas celulares dos tecidos, é necessário um microscópio, pois as células e seus componentes são tipicamente muito pequenos. Os microscópios atuais fornecem ao observador uma resolução aprimorada (capacidade de observar dois objetos próximos como objetos distintos), contraste (capacidade de detectar diferentes regiões da amostra com base na intensidade ou cor) e ampliação (capacidade de tornar pequenos objetos visíveis). Para fins comparativos, o olho humano pode resolver objetos da ordem de 0,1 milímetros, enquanto um simples microscópio de luz pode resolver objetos da ordem de 0,0002 milímetros (200 nanômetros) com uma ampliação de 1000X. Os microscópios de luz são chamados assim porque utilizam um conjunto de lentes para a ampliação da imagem transpassada por um feixe de luz.

Além do microscópio, é necessário ter uma amostra do material biológico, que pode ser obtido através de biópsias e peças cirúrgicas. Mas o maior desafio para a realização deste tipo de exame é a preservação dos tecidos, visto que o patologista deve receber o material em excelente estado de conservação. Por isso existem várias etapas de preparação com este objetivo. O processo completo pode ser definido da seguinte forma: coleta da amostra, fixação, desidratação, clarificação, inclusão (uso de parafina), microtomia (corte dos tecidos) e coloração.

Ao se remover qualquer material (órgão ou tecido) de um organismo, é necessária a realização do processo de fixação utilizando soluções químicas que tem por objetivo interromper os fenômenos de putrefação pós-morte. Este processo mantém a integridade do tecido obtido deixando-o mais parecido possível quando estava no paciente (NETO, 2012). O formalina tamponada e a solução de Bouin são exemplos de agentes fixadores que podem ser utilizados neste procedimento.

Após essa etapa, realiza-se a aplicação de soluções alcoólicas no processo de desidratação que consiste na remoção da água dos tecidos. A partir desse passo, a clarificação é realizada para a remoção do álcool, utilizando geralmente um solvente orgânico como o xilol, tornando o material mais claro ou transparente (SILVA; CORREIA, 2021).

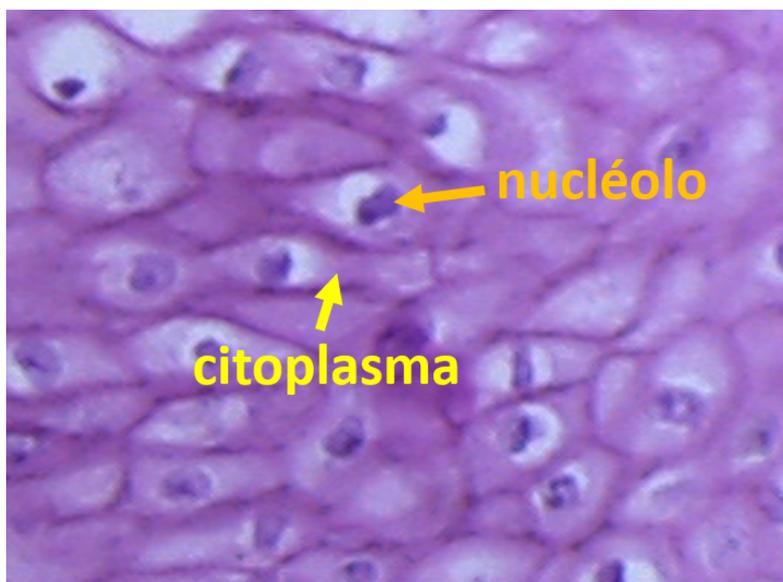
Mesmo após as etapas descritas anteriormente, as amostras ainda apresentam suas estruturas frágeis. Logo, Silva e Correia (2021) explicam que os materiais devem ser colocados em parafina derretida, processo chamado de inclusão. Os blocos de parafina que

contêm os tecidos são levados a um micrótomo¹, onde são seccionados por uma lâmina de aço ou de vidro, sendo guardados para uso posterior.

Após a aplicação da parafina, as seções ainda continuam incolores, logo, não são adequadas para o exame em microscópio. Neste caso a parafina é removida para o processo de coloração. Apesar de existirem outras formas de coloração, este trabalho se limitará a comentar o processo envolvendo as soluções hematoxilina e eosina. A primeira solução possui a natureza básica e tingem componentes ácidos de roxo-azulado, por exemplo, os nucléolos das células. Em contrapartida, a eosina possui a natureza ácida, conferindo uma coloração rósea avermelhada para componentes básicos, por exemplo, ao citoplasma das células conforme pode ser visto na Figura 2.

Após essa etapa, as lâminas podem ser analisadas através do microscópio (NUNES; CINSA, 2016). Na análise histopatológica, o patologista verifica a presença de critérios de malignidade, como o padrão de crescimento infiltrativo ou invasivo da lesão, a presença de atipias celulares (ex. elevado pleomorfismo nuclear), a presença de elevado número de células em mitose. A classificação dos tumores malignos como carcinomas escamosos assenta na identificação das características antes mencionadas (células escamosas, com pontes intercelulares - desmossomas - proeminentes e formação de pérolas de queratina), que são também visualizadas no microscópio (Figura 3).

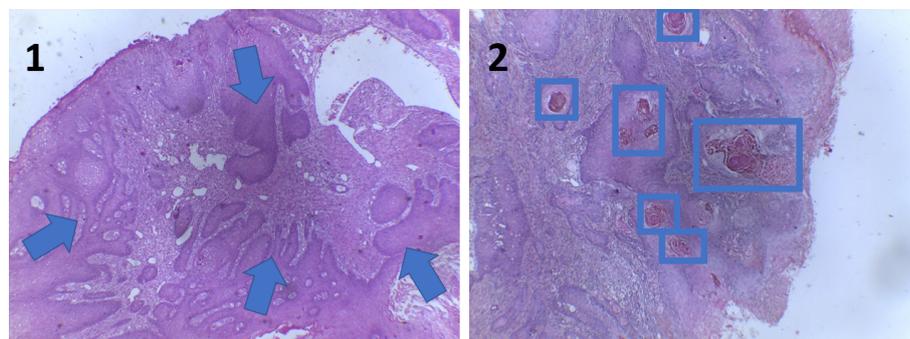
Figura 2 – Exemplo da ação da coloração com hematoxilina e eosina nos componentes do tecido peniano de lâmina histopatológica.



Fonte: elaborado pelo autor

¹ Aparelho utilizado para fazer cortes microscópicos em blocos de parafina.

Figura 3 – No primeiro exemplo, as áreas de crescimento infiltrativo (invasivo) da lesão estão indicadas conforme as setas que descrevem sua direção. No segundo exemplo, são observadas as pérolas de queratina indicadas pelos retângulos.



Fonte: elaborado pelo autor

3.3 Histopatologia Digital

Antes de discutir sobre tecnologias de reconhecimento de imagens digitais utilizando Aprendizado Profundo, esta seção traz uma breve descrição da patologia digital, o elemento básico que permite a histopatologia assistida por computador ser útil. Conforme [Al-Janabi, Huisman e Diest \(2012\)](#), a parte central da patologia é a interpretação das imagens de tecidos e suas respectivas composições celulares em uma resolução maior do que os olhos humanos são capazes de ver. Por muito tempo, o microscópio tem sido a única ferramenta disponível para este fim, sendo ao longo do tempo aperfeiçoado para fornecer imagens mais adequadas e em melhores resoluções. Mas desde a criação das câmeras digitais, que fizeram os modelos de câmeras baseados em filmes desaparecerem do mercado em pouco tempo, pode-se verificar que a histopatologia, uma subárea médica totalmente baseada em imagens, está sendo dominada por aplicações digitais, facilitando o processo de diagnóstico ([CHAN; MARINKOVICH, 2021](#)).

Inicialmente, a patologia óptica foi gradualmente mudando com a introdução de câmeras digitais que produzem imagens estáticas e câmeras de vídeo montadas em microscópio que permitem o exame ao vivo de slides utilizando *softwares* específicos para a análise. Com o passar do tempo, melhorias dessas técnicas resultaram na criação de *scanners* de slides digitais ([AL-JANABI; HUISMAN; DIEST, 2012](#)). Isso fez com que a patologia digital cada vez mais ganhasse importância e notoriedade na prática clínica moderna, sendo mais uma exigência tecnológica dentro do ambiente laboratorial. Com a evolução tecnológica, como o aumento do poder computacional, redes de comunicação mais velozes e armazenamento de dados cada vez mais baratos, trouxe a oportunidade para que os patologistas gerenciassem imagens de slides digitais com mais facilidade e flexibilidade do que podiam no passado. Permite a eles o compartilhamento de imagens para tele-patologia ([FARAHANI; PANTANOWITZ, 2015](#)) e uso clínico.

A utilização de *scanners* digitais permitiram a geração de imagens de slides inteiros. O escaneamento do tecido tingido em um lâmina de vidro produz um arquivo de imagem digital chamado de *whole slide image* (WSI). A tecnologia evoluída permitiu serem armazenadas e manipuladas em alta resolução (NIAZI; PARWANI; GURCAN, 2019). Sabendo que a qualidade desses WSIs são de total importância para fins de diagnóstico, a *Food and Drug Administration* dos Estados Unidos da América participa do processo de aprovação de *scanners* comerciais, garantindo a qualidade das imagens (HUSS; GRUNKIN, 2022).

Além disso, Chan e Marinkovich (2021) informam que vários estudos foram realizados e confirmaram que os diagnósticos utilizando imagens digitais revelaram pouca ou nenhuma diferença em relação aos exames convencionais através de microscópio. Estas imagens digitalizadas podem então ser enviadas para patologistas para análise. O fácil transporte de uma imagem digital pela internet, portanto, facilita o processo de diagnóstico, especialmente quando patologistas estão em locais distantes, sendo necessários em casos difíceis ou quando procurados para uma segunda opinião de diagnóstico. Por fim, a recente maturidade no uso da patologia digital abre um caminho oportuno para o desenvolvimento de ferramentas de diagnóstico de exames histopatológicos auxiliado por computador.

3.4 Análise de Imagens Médicas Auxiliadas por Computador

De acordo com Shen, Wu e Suk (2017), ao longo das últimas décadas, técnicas médicas de imagem, como óptica, tomografia computadorizada (TC), ressonância magnética (RM), tomografia por emissão de pósitrons (*Positron Emission Tomography*, PET), mamografia, ultrassom e raio-X, têm sido utilizadas para a detecção precoce, diagnóstico e tratamento de diversas doenças. No contexto clínico, a interpretação da imagem médica tem sido realizada principalmente por especialistas como médicos, radiologistas e patologistas.

Na histopatologia, os profissionais patologistas desempenham um papel fundamental na detecção e tratamento do câncer. A análise patológica é inerentemente subjetiva por natureza, logo, diferenças na percepção visual e no treinamento clínico podem levar a inconsistências nas opiniões diagnósticas e nos prognósticos (ESTEVA et al., 2021). No entanto, ferramentas de Inteligência Artificial (IA), como as baseadas em Aprendizado Profundo, podem fornecer suporte às tarefas médicas críticas, incluindo diagnóstico, prognóstico, segmentação patológica, monitoramento de doenças e assim por diante. Conforme Chollet (2021), a Inteligência Artificial pode ser definida como um campo de estudo da Ciência da Computação e tem como definição concisa: "O esforço para automatizar tarefas intelectuais normalmente realizadas por humanos". Entre as ferramentas automatizadas que podem ser desenvolvidas através de tecnologias de IA estão a detecção e segmentação de achados patológicos, a quantificação da extensão da doença, a

caracterização de patologias (por exemplo, benigna versus maligna ou a análise do estágio da doença) e ferramentas de software variadas que podem ser amplamente caracterizadas como suporte à decisão (ZHOU et al., 2021) (HUSS; GRUNKIN, 2022).

3.5 Redes Neurais e Aprendizado Profundo

Para que seja possível entender o conceito de Aprendizado Profundo, os algoritmos e técnicas relacionadas, esta seção discorre, primeiramente, sobre as redes neurais artificiais das quais os algoritmos de Aprendizado Profundo são baseados.

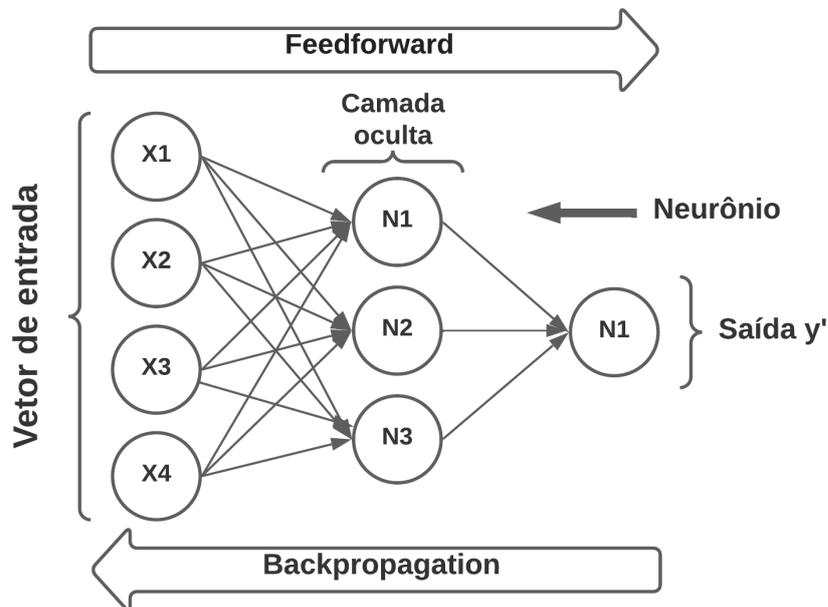
As redes neurais artificiais são algoritmos de Aprendizado de Máquina inspiradas nos neurônios biológicos do cérebro humano que são ativados sob certas circunstâncias, resultando em uma ação relacionada realizada pelo corpo em resposta a um estímulo. O Aprendizado de Máquina é uma subárea da Inteligência Artificial que, de acordo com Ketkar (2017), concentra-se no desenvolvimento de algoritmos que podem aprender como os humanos fazem, isto é, eles ficam melhores em uma tarefa ao longo do tempo, com experiência. O conceito de ativação de neurônios artificiais e a possibilidade desses neurônios aprenderem foram difundidos pelos trabalhos de Mcculloch e Pitts (1943) e Hebb (1949) na década de 1940. Mas só em 1958 com o trabalho de Rosenblatt (1958), estas ideias entraram em prática com o algoritmo perceptron, uma rede neural artificial simples que podia aprender através de um processo de atualização de parâmetros similar ao proposto por Hebb (1949).

As redes neurais artificiais (Figura 4) consistem em várias camadas de neurônios artificiais interconectados, alimentados por funções de ativação que ajudam a ligá-los ou desligá-los. Assim como os algoritmos de máquinas tradicionais, há também certos valores que as redes neurais aprendem na fase de treinamento. Cada neurônio o_i de uma determinada camada recebe os valores x_k provenientes da camada de entrada multiplicados por pesos aleatórios w_k que são, então, somados a um viés b_i exclusivo de cada neurônio (vide Equação 3.1). O resultado é então passado para uma função de ativação que decide o valor final de saída do neurônio. Neste caso, i corresponde à quantidade de neurônios da camada e k corresponde a cada entrada/peso até N.

Existem várias funções de ativação disponíveis de acordo com a natureza dos valores de entrada, como a função identidade (Equação 3.2), tangente hiperbólica (Equação 3.3), em que o resultado produzido varia de -1 a 1, e a sigmoide (Equação 3.4), em que o resultado produzido varia de 0 a 1, por exemplo.

$$o_i = \sum_{k=1}^N x_k * w_k + b_i \quad (3.1)$$

Figura 4 – Exemplo de uma rede neural multicamadas também conhecida como *multi-layer perceptron*.



Fonte: elaborado pelo autor

$$f(x) = x \quad (3.2)$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (3.3)$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.4)$$

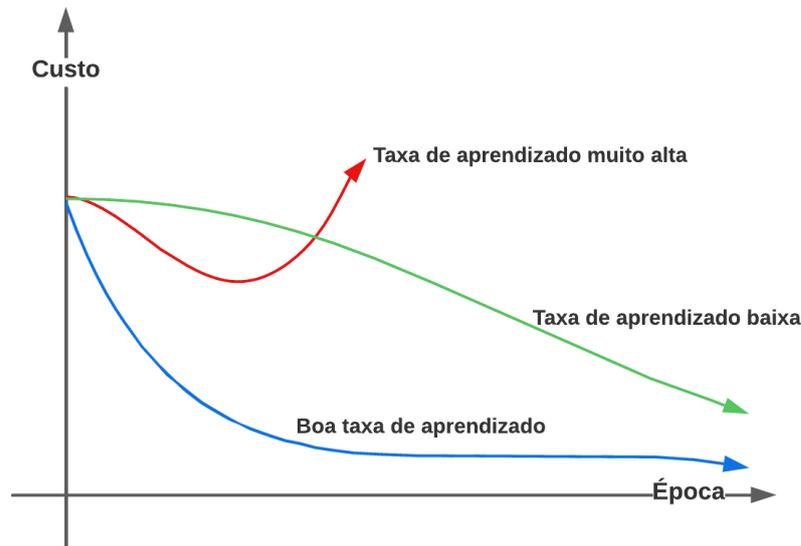
Uma vez que o valor predito é gerado pela camada final neste processo *feedforward*, a função de custo (ou *loss*) é calculada utilizando esta saída y' e o valor esperado y , a partir daí o processo de *backpropagation* é iniciado. Conforme [Matignon \(2005\)](#) e com base na Equação 3.5, a atualização dos pesos e vieses começando da última camada em direção à inicial consiste em calcular as derivadas parciais (gradiente) da função de custo C em relação a cada peso/viés da coleção θ . O sinal negativo indica a ida em direção contrária ao do gradiente. O parâmetro η representa a taxa de aprendizado da rede neural, controlando o tamanho do passo tomado para a correção do valor do peso/viés.

$$\theta_{novo} = \theta - \eta \nabla C \quad (3.5)$$

Neste caso deve-se ter cuidado ao ajustar a taxa de aprendizado, pois valores altos podem deixar o processo instável e valores muito baixos podem fazer com que o treinamento demore para convergir, conforme pode ser visto na Figura 5 ([RODRIGUES,](#)

2021). O foco geral, então, é encontrar valores ideais para os pesos neste processo de otimização que pode ser repetido várias vezes de acordo com um parâmetro que define o número de épocas.

Figura 5 – Exemplo da influência da taxa de aprendizado na convergência da rede.



Fonte: elaborado pelo autor

Pode-se ver na Figura 4 que a camada de entrada representa dimensões do vetor de entrada. Já a camada oculta representa os nós intermediários que recebem o conjunto de entradas ponderadas e produz a saída através de uma função de ativação. Por fim, a camada de saída representa a saída da rede neural e, dependendo do problema a ser resolvido, pode ter mais de um neurônio implicando uma forma diferente de calcular a função de custo. Por exemplo, para uma regressão simples (predição de um valor numérico), na camada de saída, uma função ativação do tipo identidade geraria o resultado e a função *Mean Squared Error* (MSE), representada pela Equação 3.6, calcularia o custo entre o valor esperado y e o predito y' . Já para um problema de classificação binária ou multi-classe, pode-se usar na camada de saída a função de ativação Softmax (Equação 3.7) e a função de entropia cruzada (Equação 3.8) para calcular o custo das probabilidades de cada classe.

A função de custo MSE leva em consideração a média dos quadrados das diferenças entre os valores reais e os valores preditos. Neste caso, o quadrado do erro penaliza de maneira expressiva os erros maiores. Isso significa que o aprendizado do modelo será focado em minimizar os erros mais significativos.

$$C = \frac{1}{n} \sum_{i=1}^n (y - y')^2 \quad \text{sendo } n \text{ o número de amostras} \quad (3.6)$$

Conforme [Ketkar \(2017\)](#), em um problema de classificação binária ou multi-classe, a função Softmax normaliza as saídas da camada anterior para que a soma seja igual a 1. Normalmente, as unidades da camada anterior modelam uma pontuação não normalizada da probabilidade da entrada pertencer a uma classe específica. A camada Softmax normaliza isso para que a saída represente a probabilidade para cada classe.

$$f(x)_i = \frac{e^{x_i}}{\sum_{i=1}^n e^{x_i}} \quad \text{para a classe } i = 1, 2, \dots, n \quad (3.7)$$

De acordo com [Géron \(2021\)](#), o objetivo, no caso, de uma classificação binária ou multi-classe é "ter um modelo que estime uma probabilidade alta para a classe-alvo (e consequentemente uma probabilidade baixa para as outras classes)". Para alcançar esse objetivo, é necessário minimizar a função de custo, denominada entropia cruzada (Equação 3.8), que penaliza o modelo $f(x_i)$ quando ele estima uma probabilidade baixa para uma classe-alvo y .

$$C = - \sum_{i=1}^n y_i \log f(x_i) \quad \text{para a classe } i = 1, 2, \dots, n \quad (3.8)$$

Por fim, com o progresso tecnológico, as redes neurais se tornaram mais populares, possibilitando que pesquisadores pudessem desenvolver novas arquiteturas mais poderosas. Em meados dos anos 2000, o termo Aprendizado Profundo (*deep learning*) começou a ser difundido ([HINTON; SALAKHUTDINOV, 2006](#)). De acordo com [Raschka e Mirjalili \(2019\)](#), o Aprendizado Profundo pode ser definido como “um subcampo do Aprendizado de Máquina que se preocupa em treinar a rede neural artificial com muitas camadas de forma eficiente”. Da mesma forma, [Chollet \(2021\)](#) definiu o Aprendizado Profundo como “um subcampo específico do Aprendizado de Máquina, uma nova visão sobre o aprendizado de representações a partir de dados que enfatiza o aprendizado de sucessivas “camadas” de representações cada vez mais significativas”.

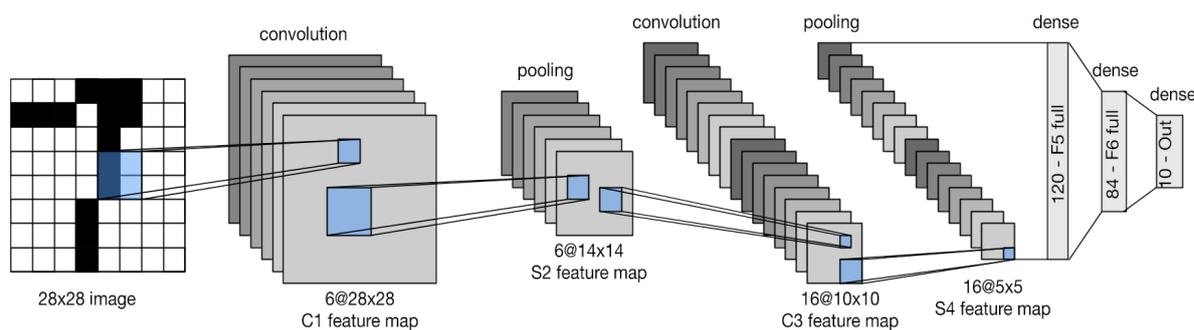
Existem várias redes neurais que estão incluídas neste subcampo, como as redes multi-layer perceptron, redes convolucionais, redes recorrentes, as redes *Long Short-Term Memory* (LSTM), entre outras. Sendo as redes convolucionais, especificamente as DenseNets, as utilizadas neste trabalho.

3.5.1 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (ou *Convolutional Neural Networks* - CNNs) são arquiteturas especiais inspiradas nos mecanismos biológicos da visão dos seres vivos, sendo utilizadas, por exemplo, em tarefas de classificação, segmentação e reconhecimento de objetos em imagens digitais. O que as tornam diferentes das demais arquiteturas baseadas em rede neural é a utilização de operações de convolução em algumas de suas camadas,

mas de forma complementar, sua estrutura normalmente é composta por outros tipos de camadas, como as camadas de *pooling* e a totalmente conectada. A primeira publicação desse tipo de rede neural foi realizada por [LeCun et al. \(1989\)](#) na qual foi apresentada a LeNet (Figura 6). Mas elas só se tornaram populares através do trabalho de [Krizhevsky, Sutskever e Hinton \(2012\)](#) para o desafio ImageNet de 2012, como comentado em [Zhang et al. \(2021\)](#). A partir disso, muitas outras novas arquiteturas foram surgindo, sendo bastante aplicadas no contexto de análise de imagens médicas.

Figura 6 – Rede convolucional clássica - LeNet5.



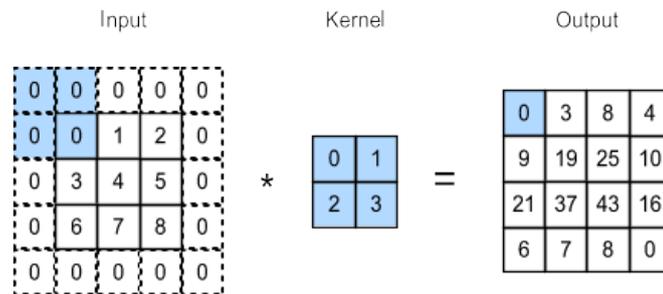
Fonte: ([ZHANG et al., 2021](#))

Conforme pode ser visto nesta simples rede CNN da Figura 6, os resultados das camadas convolucionais, a partir das imagens de entrada, são chamados de mapas de características (*feature maps*). Estas servem de entrada para a camada seguinte até se transformarem em um vetor de características utilizado na primeira camada densa ou totalmente conectada. Conforme [Mishra \(2020\)](#), o processo de convolução funciona da seguinte forma: uma pequena matriz chamada de *kernel* com determinados elementos é utilizada sobreposta à imagem, operações de multiplicação e soma são efetuadas; e em um movimento chamado de deslizamento de janela, essas operações vão sendo realizadas conforme um número determinado de passos (*stride*) até que uma outra matriz com os valores resultantes seja construída (vide Figura 7). O ponto chave das redes convolucionais é que nessas camadas os pesos a serem aprendidos são exatamente os valores desses *kernels*.

Já a camada *pooling* tem como objetivo reduzir a resolução desses mapas de características e deixar a rede menos sensível a pequenas distorções, ou seja, mais invariante, além de reduzir o *overfit* ([GRZYBOWSKI, 2021](#)). A operação de *pooling*, também conhecida como *subsampling* (subamostragem), funciona da seguinte forma: através da técnica de deslizamento de janela, uma área de tamanho, por exemplo, 2x2, é sobreposta em uma matriz de entrada; operações nos pixels desta vizinhança baseado ou na média ou no valor máximo ou no valor mínimo resultam em valores que farão parte de uma matriz resultante (vide Figura 8).

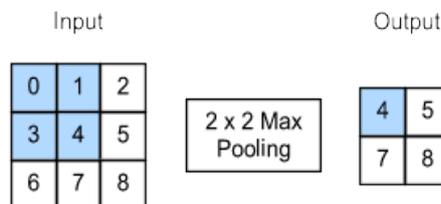
Por fim, as camadas densas (*dense layers*) ou totalmente conectadas (*fully connected*

Figura 7 – Exemplo da técnica de convolução: neste caso, para preservar o tamanho da matriz, utilizou-se a adição de bordas (*padding*) com os valores iguais a zero.



Fonte: (ZHANG et al., 2021)

Figura 8 – Exemplo da técnica de *pooling*, neste caso, o *max pooling* (valor máximo).



Fonte: (ZHANG et al., 2021)

layers) trabalham exatamente como uma rede neural padrão, na qual todos os neurônios de uma certa camada são interligados a todos os neurônios da camada anterior. Geralmente, a maioria das camadas de uma rede convolucional utiliza em seus neurônios a função de ativação *Rectified linear unit* (Relu) ou suas derivadas:

$$\text{relu}(z) = \max(0, z) \quad (3.9)$$

3.5.2 O Problema do Vanishing Gradient nas Redes Neurais Profundas

De acordo com Erdmann et al. (2021), conforme as redes neurais aumentam de profundidade, ou seja, ganham novas camadas, mais parâmetros precisam ser aprendidos. Por conseguinte, aumenta-se a probabilidade de ocorrer o esvanecimento dos gradientes (*vanishing gradient*) no processo de *backpropagation*. Certas funções de ativação, como a função sigmoide, transformam um espaço de entrada em um pequeno espaço (entre 0 e 1) e, neste caso, se uma grande mudança na entrada da função sigmoide ocorrer, uma pequena mudança será feita na saída. Logo, a derivada se torna pequena.

Pela regra da cadeia, as derivadas de cada camada são calculadas da camada final para a inicial da rede com o objetivo de atualizar os pesos e vieses de todas as camadas. No entanto, quando n camadas ocultas usam uma função de ativação como a função

sigmoide, as pequenas derivadas são multiplicadas juntas. Assim, o gradiente diminui exponencialmente à medida que o erro é propagado até as camadas iniciais.

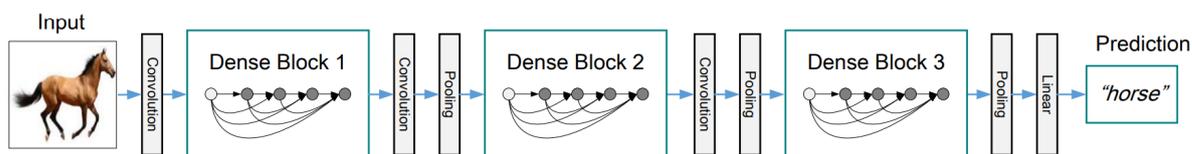
Um pequeno gradiente significa que os elementos a serem aprendidos nas camadas iniciais não serão atualizados efetivamente a cada iteração de treinamento. Uma vez que essas camadas são muitas das vezes importantes para o reconhecimento dos elementos centrais dos dados de entrada, isso pode levar a um modelo preditivo impreciso.

Uma forma de diminuir esse problema pode ser utilizando a função de ativação Relu que não produz uma pequena derivada. Como a maioria das arquiteturas de rede neurais de última geração são muitas vezes profundas e amplas, a Relu é a função de ativação mais utilizada hoje em dia (TAN; LIM, 2019). Outra forma é a utilização de conexões de atalho com o objetivo de fortalecer o *backpropagation*. Nesse caso, têm-se como exemplo as conexões residuais (HE et al., 2015) e as densas (HUANG et al., 2017).

3.5.3 Arquitetura DenseNet

As DenseNets (Figura 9) foram apresentadas em Huang et al. (2017). Estas tem como principais objetivos: reduzir o efeito chamado de *vanishing gradient*, que ocorre com frequência em redes neurais profundas, levando a rede a parar de aprender em um processo de treinamento, além disso, fortalecem a propagação das características e reduzem de forma bastante significativa a quantidade de parâmetros a serem aprendidos.

Figura 9 – Exemplo de uma DenseNet com três blocos densos.

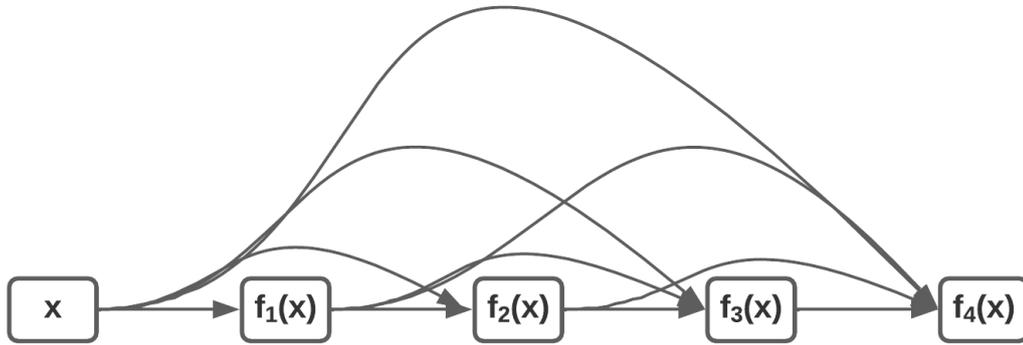


Fonte: (HUANG et al., 2017)

A principal característica desse tipo de CNN é a forma como os dados são enviados entre as camadas. Cada uma delas possui conexões (ou atalhos) com as outras camadas posteriores utilizando a operação de concatenação (vide Figura 10), algo bem semelhante às ResNets (HE et al., 2015), apesar destas usarem um tipo de operação (soma) diferente entre as camadas. Estes atalhos fazem com que as camadas recebam um conhecimento coletivo das demais.

Uma rede DenseNet é constituída por vários blocos densos compostos por operações convolucionais contendo os atalhos comentados anteriormente. Cada bloco é separado por uma camada de transição composta por uma camada de convolução e outra de *pooling*, que tem por objetivo reduzir a resolução dos mapas de características (*feature maps*)

Figura 10 – Exemplo de uma *dense block* e as conexões entre as camadas.



Fonte: elaborado pelo autor

gerados que serão enviados para as próximas camadas. Isso é necessário pois as camadas internas dos blocos preservam a resolução dos *feature maps*. Existem várias versões (Figura 11) de DenseNets provenientes do trabalho original (HUANG et al., 2017), numeradas pela quantidade de camadas. Neste caso, a arquitetura utilizada neste trabalho possui exatamente por padrão 201 camadas, sua escolha se justifica por ter tido um desempenho superior em relação aos outras versões da DenseNet em experimentos realizados.

Figura 11 – Tabela proveniente do trabalho original de Huang et al. (2017) onde podem ser vistas as arquiteturas e suas configurações de camadas.

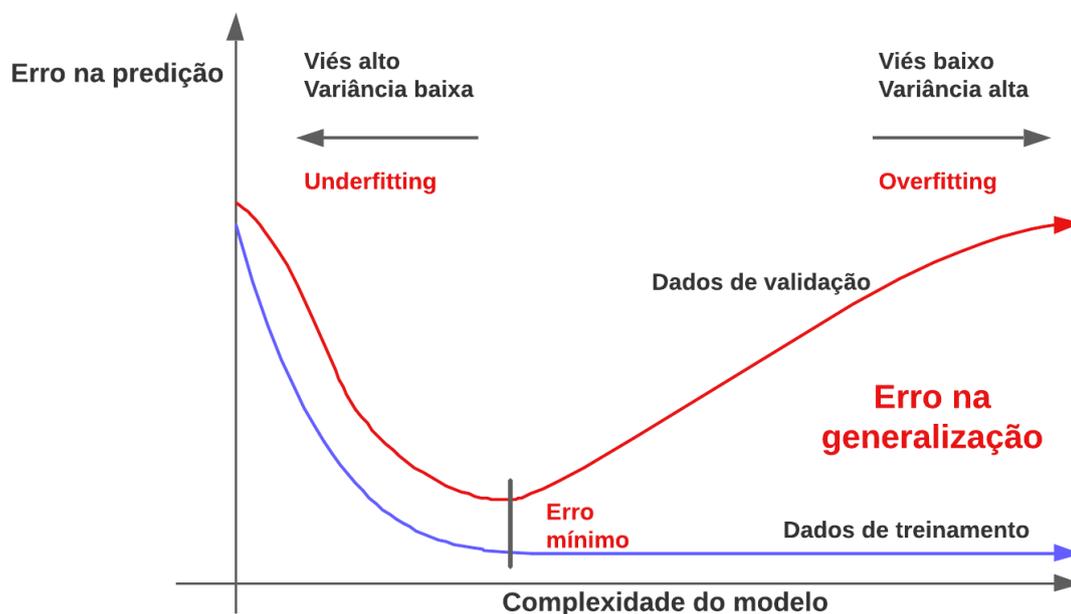
Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

Fonte: (HUANG et al., 2017)

3.5.4 Overfitting e Underfitting

O *overfitting* e o *underfitting* são problemas que podem ocorrer no Aprendizado de Máquina supervisionado, o que inclui as redes neurais em geral. O *overfitting* é uma diferença muito significativa entre o resultado da avaliação do modelo nos dados de treinamento e nos que foram separados para teste ou validação, ou seja, quer dizer que o modelo não generalizou. Conforme Shorten e Khoshgoftaar (2019), esse problema refere-se ao fenômeno ocorrido quando o algoritmo de Aprendizado de Máquina aprende uma função com variância muito alta, como modelar perfeitamente os dados de treinamento. Para evitar este problema há necessidade de avaliar o modelo gerado, efetuar ajustes nos parâmetros dependendo do algoritmo de aprendizagem de máquina utilizado ou aplicar alguma técnica apropriada para o contexto. Já o *underfitting* acontece quando o modelo não aprendeu o suficiente para reconhecer os padrões que o farão distinguir as amostras dentre determinadas classes, por exemplo, significando também que o modelo possui um erro alto (alto viés) tanto no treinamento como no teste e validação. Na Figura 12 podem ser observados estes fenômenos, no qual o ponto de erro mínimo indicado é o ideal para o modelo generalizar bem.

Figura 12 – Nesta figura é exemplificada a ocorrência do *overfitting* e *underfitting*.



Fonte: elaborado pelo autor

Como forma de reduzir o overfit, alguns métodos de regularização de redes neurais podem ser empregados. Dentre eles, têm-se a Transferência de Aprendizado e *Data Augmentation*, ambos são usados neste estudo e são descritos nas próximas seções.

3.6 Transferência de Aprendizado e Fine-tuning

Uma das desvantagens do uso do Aprendizado Profundo está no fato de que os algoritmos de aprendizado dependem de uma quantidade extremamente grande de amostras para treinamento, entre milhares e milhões, por exemplo. Caso contrário, a probabilidade de *overfitting* é muito grande. Por isso, as tarefas de Aprendizado de Máquina que envolvem imagens médicas comumente são um desafio para o cientista da computação, pois adquirir e rotular uma quantidade imensa de imagens são atividades árduas e difíceis, visto que a obtenção de imagens médicas envolve vários fatores como a disponibilidade do paciente, do médico, de equipamentos, além de tramites legais e éticos dependendo de cada país (MORID; BORJALI; Del Fiol, 2021).

Como forma de evitar ou minimizar o problema do *overfitting* existem várias técnicas que podem ser empregadas, como o *dropout*, *batch normalization*, *transferência de aprendizado*, *pretraining*, funções de regularização e *data augmentation*. O uso da técnica de transferência de aprendizado é uma das mais populares. Nesta abordagem os pesos aprendidos por uma rede neural em um determinado domínio de problema podem ser reutilizados em um outro domínio diferente, mas, como a camada de classificação é específica para o problema original, faz-se necessária a alteração da última camada para adaptar esta arquitetura de CNN a um novo problema. Conforme Tajbakhsh et al. (2016), as camadas iniciais das redes convolucionais aprendem características de baixo nível que são aplicáveis de maneira geral à maioria das tarefas de visão computacional. Já as provenientes de camadas mais profundas aprendem características de alto nível específicas ao domínio do problema ao qual elas estão sendo aplicadas. A partir dessas afirmações, uma técnica auxiliar pode ser empregada juntamente com a transferência de aprendizado com o objetivo de melhorar os resultados do treinamento, chamada de *fine-tuning*, que consiste em treinar as camadas agregadas e também aquelas provenientes da CNN pré-treinada.

3.7 Data Augmentation

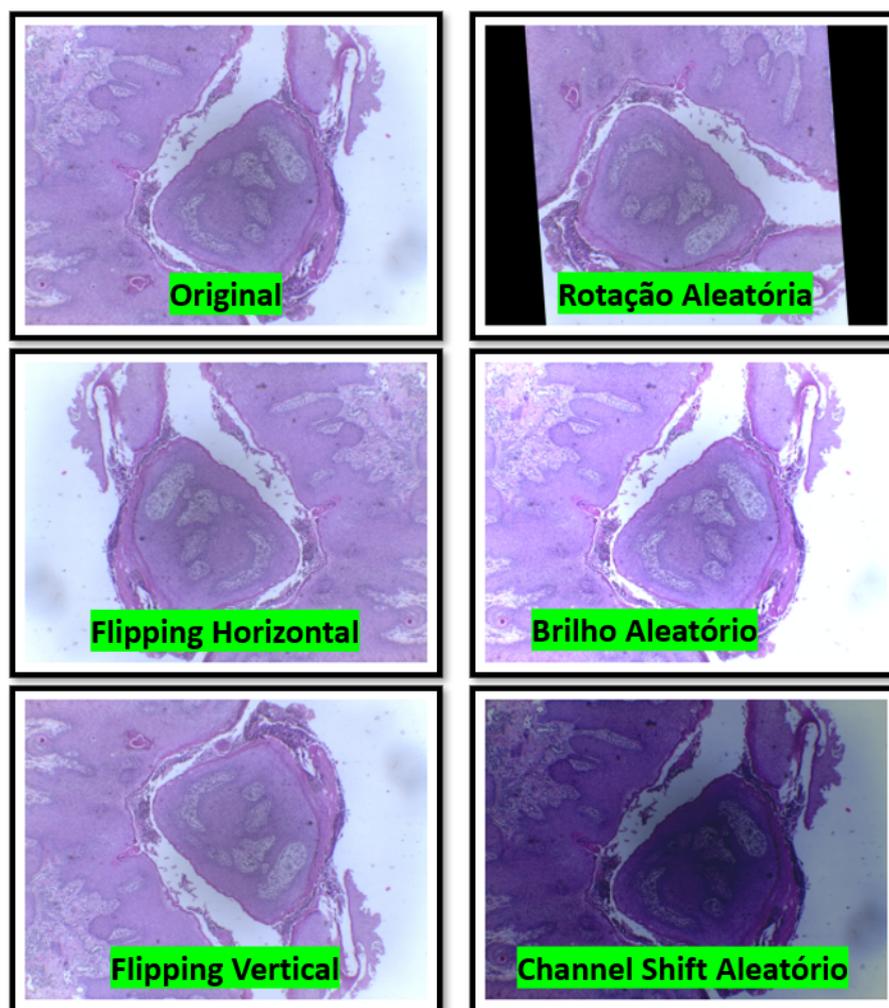
O *data augmentation* engloba um conjunto de técnicas que aprimoram o tamanho e a qualidade dos conjuntos de dados de treinamento para que melhores modelos de Aprendizado Profundo possam ser construídos (SHORTEN; KHOSHGOFTAAR, 2019). Em primeiro lugar, o reconhecimento de imagens é uma tarefa difícil e em exemplos discriminatórios clássicos como gato versus cão, o software de reconhecimento de imagens deve superar questões como a iluminação, oclusão, fundo, escala, dentre outras. A tarefa do *data augmentation* é inserir essas invariâncias no conjunto de dados de treinamento de modo que os modelos resultantes tenham um bom desempenho no conjunto de testes.

As formas mais simples de *data augmentation* para imagens são representadas pelas

seguintes transformações: rotação; inversão (*flipping*) horizontal e vertical; brilho aleatório; e o *channel shift* aleatório que desloca aleatoriamente os valores em cada canal de acordo com um determinado valor (vide Figura 13). Apesar disso, existem outras formas como filtros de *kernel*, mistura de imagens, *random erasing*, alterações no espaço de cor, uso de *Generative Adversarial Networks* (GANs) e transferência de estilo via redes neurais, dentre outras.

O uso de certas transformações ou algoritmos depende muito do problema a ser aplicado, pois o cientista da computação deve conhecer o domínio do problema, evitando, assim, que transformações prejudiquem a rotulagem da imagem. Por exemplo, para imagens que representam números ou letras, não podemos usar de forma indiscriminada as transformações de rotação ou *flipping*, visto que uma imagem que representa o número 6 pode acabar se transformando em uma imagem que representa o número 9.

Figura 13 – Exemplos de transformações aplicadas em imagens histopatológicas no processo de *data augmentation*.



Fonte: elaborado pelo autor

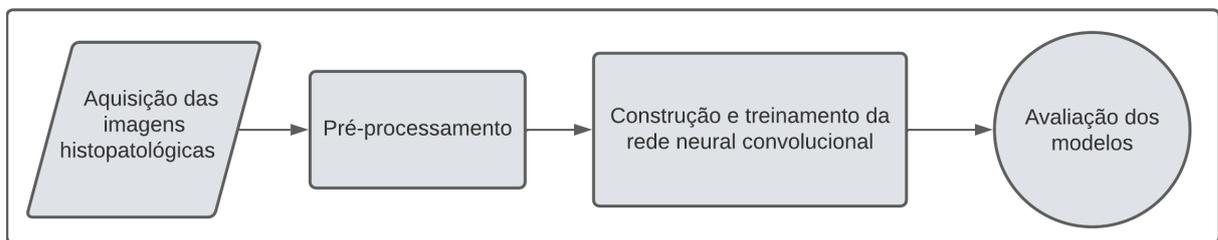
3.8 Considerações Finais

Neste capítulo foram descritos conceitos envolvendo o câncer de pênis, a análise histopatológica e as tecnologias baseadas em Aprendizado Profundo utilizadas na tarefa de reconhecimento de imagens digitais. As informações expostas e técnicas apresentadas formam uma base teórica para entendimento do método proposto (Capítulo 4) e do estudo de ablação (Seção 5.2/Capítulo 5).

4 Metodologia

Neste capítulo é detalhado o método proposto que tem as seguintes etapas: aquisição das imagens; pré-processamento das imagens através do algoritmo *Contrast Limited Adaptive Histogram Equalization* (CLAHE); construção e treinamento da rede neural convolucional e avaliação dos modelos gerados. A Figura 14 apresenta o fluxo das etapas propostas de maneira resumida.

Figura 14 – Diagrama do método proposto.



Fonte: elaborado pelo autor

4.1 Aquisição das imagens histopatológicas

A base de imagens utilizadas neste trabalho foi construída pelo Projeto Câncer de Pênis da Amazônia Legal com apoio dos docentes e alguns alunos do Núcleo de Computação Aplicada da Universidade Federal do Maranhão. Esta é composta por 194 imagens RGB de resolução de 2048x1536 pixels. Estes arquivos foram agrupados por ampliação e situação patológica conforme Tabela 2.

Tabela 2 – Distribuição das imagens conforme ampliação e classificação patológica.

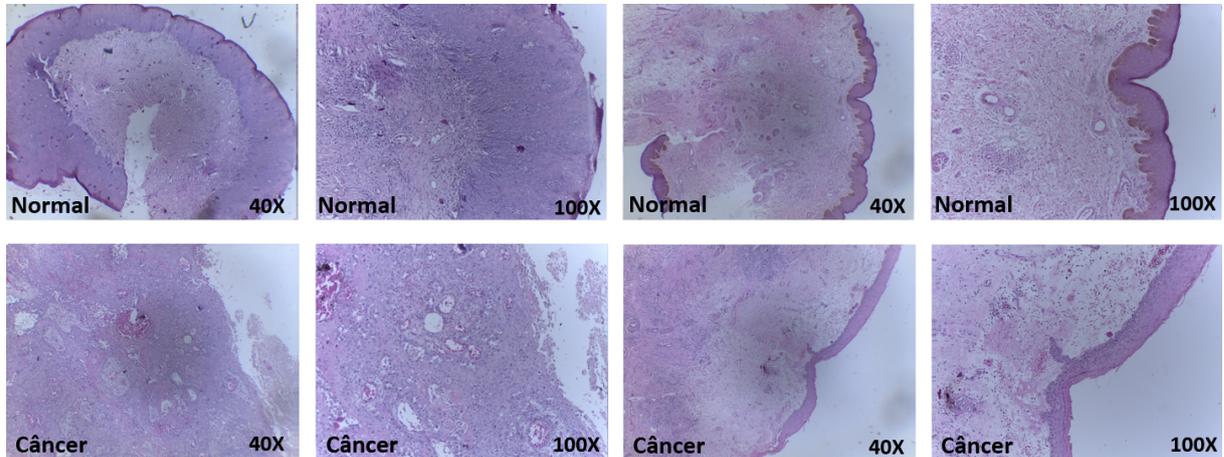
Categoria/Ampliação	40X	100X
Normal	40	40
Câncer (carcinoma de células escamosas)	57	57
Total	97	97

Fonte: elaborado pelo autor

O processo de captura das imagens foi realizado em 2021 utilizando amostras de tecidos da região do pênis, provenientes de biopsias realizadas em pacientes em tratamento no estado do Maranhão, Brasil. As amostras foram tingidas com Eosina e Hematoxilina e depois fotografadas por dois alunos de pós-graduação através de uma câmera de alta definição (Leica ICC50 HD) acoplada a um microscópio de campo claro (Leica DM500). Com auxílio de um software específico (Leica Aperio ImageScope), as imagens foram

analisadas e classificadas por dois patologistas. Alguns exemplos dessas imagens são apresentados na Figura 15.

Figura 15 – Exemplos de imagens histopatológicas de câncer de pênis por categoria e ampliação.



Fonte: elaborado pelo autor

4.2 Pré-processamento

Conforme Pal e Sudeep (2016), os dados brutos, se aplicados a qualquer método de classificação, não produzem uma boa precisão na maioria das vezes. Por isso é muito importante realizar o pré-processamento nas imagens antes de realizar o treinamento do modelo e a classificação das amostras. Esta etapa tem por objetivo melhorar a qualidade dos dados de entrada, diminuindo ruídos, corrigindo defeitos ou realçando algum detalhe importante, obtendo, assim, uma melhor discriminação dos objetos localizados na imagem.

Para que não houvessem problemas com o consumo de memória da Unidade de Processamento Gráfico no processo de treinamento da rede neural, as imagens foram redimensionadas de forma proporcional para resolução de 224x224 *pixels*, a mesma utilizada para as imagens da ImageNet no processo de treinamento com a DenseNet que resultou no modelo pré-treinado adotado por esse método.

Ao verificar as imagens, percebeu-se que apresentam diferenças na distribuição de iluminação. Para minimizar o efeito, usamos o algoritmo *Contrast Limited Adaptive Histogram Equalization* (CLAHE) (ZUIDERVELD, 1994). De acordo com Kumar e Shaik (2015), o CLAHE é uma evolução do *Adaptive Histogram Equalization* (AHE) e de seu precursor mais básico, o *Histogram Equalization* (HE). A principal diferença em relação aos seus precursores está na forma como é aplicada a equalização do histograma, não sendo global, mas sim em regiões de tamanho ajustável. A imagem é dividida em blocos

que são equalizados de forma adaptativa, depois são combinados e tratados para evitar qualquer efeito de borda entre eles, utilizando interpolação bilinear. Justifica-se o uso deste algoritmo, pois as formas mais simples de melhoria de contraste como o HE e AHE não deram bons resultados em comparação com o algoritmo CLAHE nos experimentos realizados pelo autor deste trabalho.

As imagens ainda foram transformadas do espaço de cor RGB para YUV (YANG; YUHUA; ZHAO GUANG, 2007). Essa transformação é necessária porque todos os canais do espaço RGB carregam informações de cor, logo, efeitos indesejáveis nas cores das imagens seriam perceptíveis após a equalização destes canais. Após o redimensionamento e transformação de cor, o algoritmo CLAHE foi aplicado no canal Y (VILLÁN, 2019). Esse canal representa as informações de intensidade. Após esta operação, as imagens equalizadas foram devolvidas para RGB. Apesar do autor deste trabalho ter experimentado outros espaços de cores, como o CIELab (L^* indica a luminosidade e o a^*b^* as coordenadas cromáticas) e HSV (*Hue, Saturation and Value*), os resultados foram ruins em relação aos resultados baseados na utilização do espaço de cor YUV.

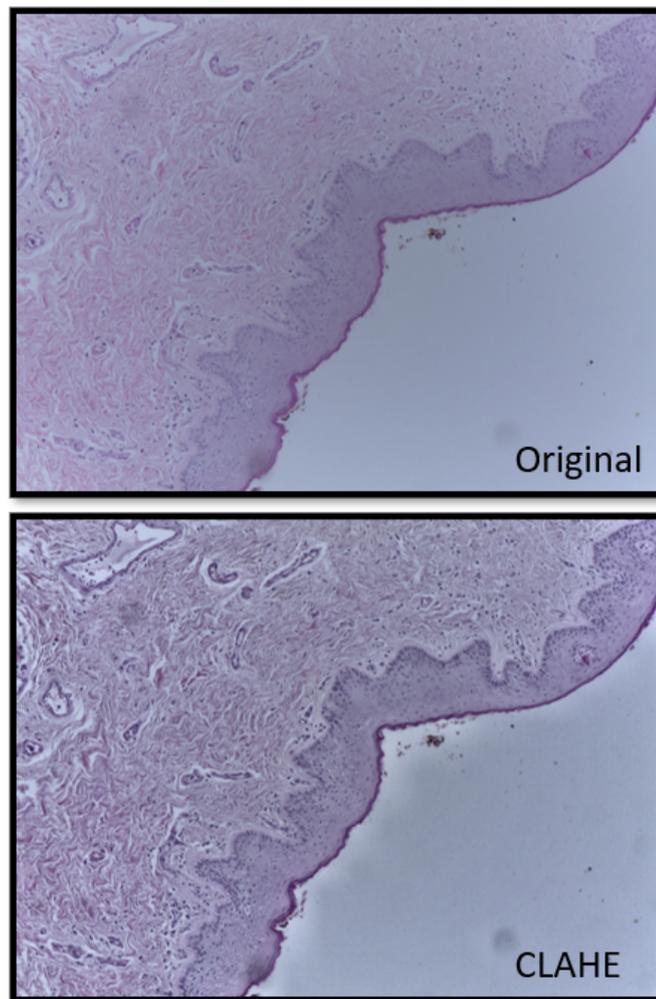
Um exemplo do pré-processamento aplicado é demonstrado na Figura 16. Neste exemplo verifica-se que as estruturas das células atingem um nível de contraste mais perceptível do que na imagem original, tornando a sua distinção em relação do fundo mais simples.

4.3 Construção e Treinamento da Rede Neural Convolutacional

O método proposto utiliza um modelo pré-treinado através da arquitetura DenseNet-201 na base de imagens ImageNet e está dividido em duas etapas de treinamento. A aplicação de transferência de aprendizado neste problema é necessária visto que a base de imagens histopatológicas possui poucas amostras. Para adaptar a rede convolutacional ao contexto do problema, as camadas de classificação utilizadas para o desafio ImageNet foram retiradas. As outras camadas da rede permaneceram congeladas, ou seja, na primeira etapa do processo de treinamento os pesos das camadas convolucionais não são atualizados. No lugar da camada de classificação removida foram adicionadas três camadas: uma camada de *average pooling*, uma *fully connected* (FC), que utiliza a função de ativação Relu com 256 neurônios, e uma camada de saída que utiliza o algoritmo Softmax. Além disso, entre as duas últimas camadas, foi adicionado um *dropout* com uma probabilidade 0,35 definida empiricamente para que, aleatoriamente, alguns neurônios sejam desativados com o intuito de reduzir o *overfitting* (HINTON et al., 2012).

Com a finalidade de especializar os pesos para o problema, foi realizado um *fine-tuning*, descongelando todas as camadas da rede, possibilitando o treinamento de todos os pesos da rede neural convolutacional. Este método, então, realiza duas etapas de treinamento,

Figura 16 – Aplicação do algoritmo de realce CLAHE em uma imagem histopatológica de câncer de pênis.

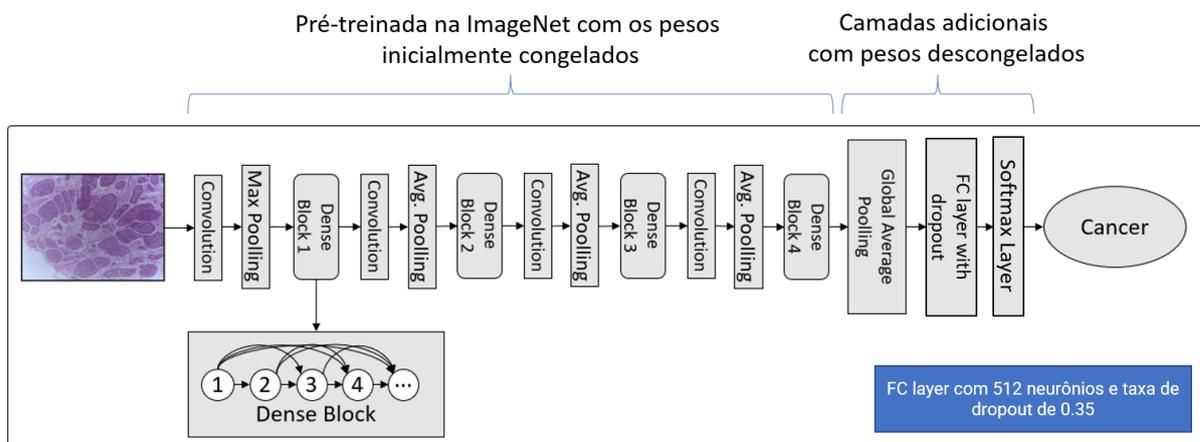


Fonte: elaborado pelo autor

primeiramente com todas as camadas convolucionais congeladas e, depois, com todas as camadas descongeladas.

Na Figura 17 é mostrada a arquitetura de CNN com os blocos densos, suas conexões entre as camadas de transição e as camadas adicionadas pelo autor que são adequadas para este problema de classificação binária. Por fim, a utilização da DenseNet-201, neste trabalho, justifica-se por ter tido um desempenho superior em experimentos realizados em relação a outras arquiteturas como a Xception (CHOLLET, 2016) e Inception ResNet V2 (SZEGEDY et al., 2016), por exemplo. Um comparativo do desempenho da DenseNet em relação às duas redes neurais citadas anteriormente foi adicionado ao Capítulo 5.

Figura 17 – Arquitetura DenseNet-201 adaptada ao problema de classificação binária.



Fonte: elaborado pelo autor

4.4 Avaliação

Após todo o processo de treinamento em duas etapas da rede DenseNet-201 na base de imagens histopatológicas de câncer de pênis, há a necessidade de testar se o modelo gerado possui um bom desempenho na classificação de novas imagens, ou seja, avaliar a generalização do modelo. Para isso utilizou-se a validação cruzada *K-fold*. Nesta técnica o conjunto de amostras é dividido em *k* subconjuntos de amostras, cada *k* rodada da execução do processo de treinamento e teste. Uma das partições sem repetição é utilizada para o teste e as outras para o treinamento, como pode ser observado na Figura 18. Assim, o teste é realizado com todo conjunto de amostras e os resultados utilizando uma métrica específica ou um conjunto delas podem ser calculados através da média, obtendo-se o desvio padrão também.

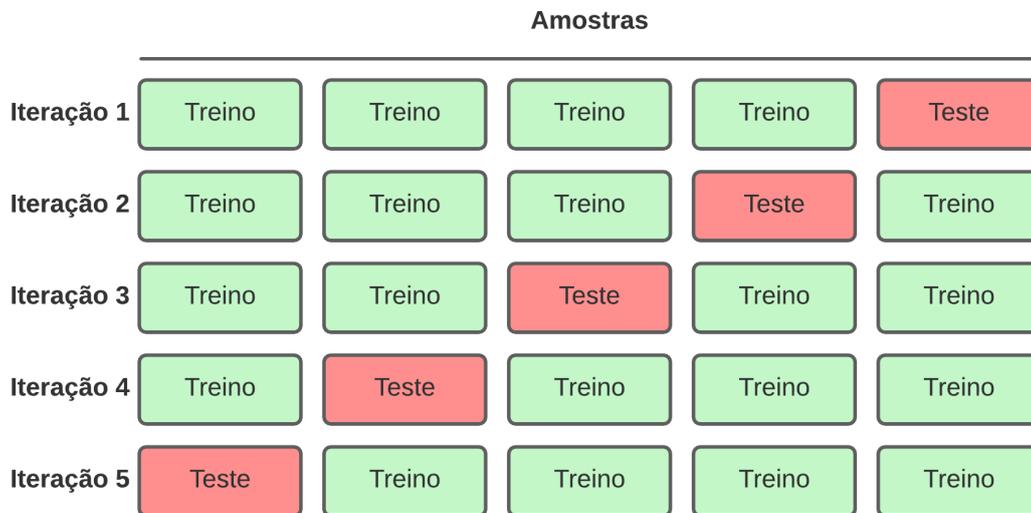
No caso da classificação binária que é o objeto deste trabalho, há a possibilidade de se obter várias métricas. Para explicar os indicadores utilizados neste estudo, a matriz de confusão observada na Tabela 3 será utilizada. Esta tabela é um resumo do desempenho de um algoritmo de classificação, dando uma visão geral do que o modelo está acertando e quais possíveis erros está cometendo.

Tabela 3 – Matriz de confusão para tarefas de classificação binária.

		Valor Real	
		Câncer	Normal
Predito	Câncer	Verdadeiro Positivo (VP)	Falso Positivo (FP)
	Normal	Falso Negativo (FN)	Verdadeiro Negativo (VN)

Fonte: elaborado pelo autor

A mais comum das métricas para modelos de classificação, a acurácia (equação

Figura 18 – Exemplo de particionamento utilizando validação cruzada com cinco *folds*.

Fonte: elaborado pelo autor

4.1), mede o número de previsões corretas como uma porcentagem do número total de previsões feitas.

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (4.1)$$

A sensibilidade ou *recall* (Equação 4.2) é a razão de observações positivas previstas corretamente para todas as observações da classe real. Essa métrica avalia a capacidade do modelo de Aprendizado de Máquina de detectar com sucesso resultados classificados como positivos. Quando o *recall* é 1.0 ou 100% conclui-se que não houve nenhum falso negativo.

$$Sensibilidade = Recall = \frac{VP}{VP + FN} \quad (4.2)$$

A especificidade (Equação 4.3) é a razão de observações negativas previstas de maneira correta para todas as observações da classe real. Ao contrário do *recall*, essa métrica avalia a capacidade do modelo de Aprendizado de Máquina de detectar com sucesso resultados classificados como negativos.

$$Especificidade = \frac{VN}{FP + VN} \quad (4.3)$$

A precisão (Equação 4.4) é a razão de observações positivas previstas corretamente para o total de observações positivas previstas, ou seja, busca-se saber qual a proporção

dos dados classificados como câncer eram realmente câncer.

$$Precisão = \frac{VP}{VP + FP} \quad (4.4)$$

O F1-Score (equação 4.5) é definido como a média harmônica da precisão e sensibilidade. Isto significa que esta métrica dá igual peso à precisão e à sensibilidade:

$$F1 - Score = 2 * \frac{Precisão * Sensibilidade}{Precisão + Sensibilidade} \quad (4.5)$$

De forma complementar, foram realizadas avaliações nas imagens de teste com o intuito de verificar quais regiões a rede neural convolucional levou em consideração para tomar determinada decisão de classificação. Isso é chamado de explicabilidade do modelo. De acordo com Borg et al. (2021), nas redes neurais convolucionais, é difícil analisar o que os nós individuais representam em relação às imagens e como eles contribuem para as previsões finais. Para apoiar a explicabilidade neste tipo de algoritmo de Aprendizagem de Máquina, utilizou-se o algoritmo *Gradient-weighted Class Activation Mapping*, conhecido como Grad-CAM (SELVARAJU et al., 2019). Este algoritmo utiliza os gradientes, especificamente, os da pontuação de classe (antes da Softmax) referentes aos mapas de ativação das camadas convolucionais. Um *global average pooling* é realizado nesses gradientes para obter pesos de importância dos neurônios. Em seguida, é calculada uma combinação linear reduzida de mapas de ativações com esses pesos, seguido da aplicação da função ReLu. Após isso, a matriz resultante é aumentada com interpolação bilinear para corresponder à dimensão da imagem de entrada. Esta abordagem permitiu produzir uma imagem do tipo mapa de calor sobreposta em cada imagem classificada pelo modelo treinado pela rede DenseNet.

4.5 Considerações Finais

Neste capítulo, foram detalhadas as etapas de aquisição do *dataset* de imagens histopatológicas de câncer pênis (Seção 4.1); o pré-processamento das imagens via algoritmo CLAHE (Seção 4.2); as definições arquiteturais de construção e treinamento da DenseNet (Seção 4.3); o processo de avaliação dos modelos gerados através da validação cruzada e a aplicação do algoritmo Grad-CAM para a explicabilidade dos modelos (Seção 4.4). No próximo capítulo, os resultados dos experimentos realizados nas partições de testes são avaliados, comparados e discutidos.

5 Resultados

Neste capítulo são detalhados os resultados dos experimentos realizados para validação do método proposto. O modelo baseado na DenseNet foi treinado utilizando uma placa de vídeo Nvidia Geforce 3060 RTX de 12GB memória, com o *dataset* de imagens histopatológicas de câncer de pênis para ambos os tipos de ampliação, em conformidade com a técnica de transferência de aprendizado. Neste caso, utilizou-se a biblioteca Keras (CHOLLET et al., 2015) e a scikit-learn (PEDREGOSA et al., 2011) para subsidiar todos os experimentos.

O primeiro treinamento (aquecimento da rede auxiliar adicionada ao final) foi realizado com 35 épocas utilizando o otimizador Adam e taxa de aprendizado de 0,0001. O *batch size* foi ajustado para 32 e alguns experimentos utilizaram de *data augmentation* com transformações (*flip horizontal*, *vertical* e rotação aleatória até 90°). Após este processo, foi realizada uma segunda etapa de treinamento realizando um *fine-tuning* no modelo resultante como forma de aprimorar os resultados. Todas as camadas desta rede foram descongeladas e o modelo foi treinado novamente com 8 épocas utilizando uma taxa de aprendizado igual a 0,00001. Todos os parâmetros descritos foram escolhidos de forma empírica pelo autor.

Os experimentos foram executados utilizando validação cruzada K-Fold e assim mensurar a generalização do modelo. Cada experimento foi configurado para 5 *folds* estratificados, sendo que para cada uma das 5 rodadas, um *fold* diferente de imagens é selecionado para o teste, os outros são utilizados para compor os particionamentos de treino e validação, respectivamente 80% e 20% também de maneira estratificada.

Com o objetivo de avaliar o método, utilizou-se as seguintes métricas apresentadas nas tabelas de resultados: acurácia (ac.), sensibilidade (sen.), especificidade (esp.), precisão (prec.) e F1-Score, sendo esta última utilizada para indicar o melhor modelo, pois agrega dois indicadores relacionados às amostras positivas (com câncer). Esses valores médios baseados nos *folds* estão acompanhados do desvio padrão que resume a variação esperada no desempenho do modelo. Todas as tabelas de resultados são agrupadas pela ampliação (amp.), uso de pré-processamento (proc.) e *data augmentation* (aug.).

5.1 Resultados dos Experimentos com o Método Proposto

Como apresentado na Tabela 4, a aplicação da operação de realce CLAHE contribuiu para que o método obtivesse um bom resultado com a ampliação de 40X tomando como base a métrica F1-Score (média harmônica entre precisão e a sensibilidade). Logo, para

experimento 3 o resultado foi $97,39 \pm 2,13$. Para a ampliação 100X verifica-se que o *data augmentation* contribuiu juntamente com o algoritmo CLAHE no experimento 8, que obteve $97,31 \pm 3,62$ de F1-Score.

Os modelos treinados ainda tiveram resultados significativos em relação à medida de sensibilidade. Esta métrica informa a proporção de imagens que possuem câncer e que foram classificadas como positivo. O experimento 3, realizado nas imagens de ampliação 40X, obteve o resultado de $98,33 \pm 3,33$. Verifica-se neste experimento a maior estabilidade com menor desvio padrão em relação outros experimentos. Para a ampliação 100X, o experimento 8 resultou em $98,18 \pm 3,64$.

Tabela 4 – Comparativo de métricas entre os diversos experimentos que comprovam o desempenho do método baseado na DenseNet.

Resultados dos Experimentos DenseNet							
Amp.	Proc.	Aug.	Ac.	Sen.	Esp.	Prec.	F1-Score
40X	Não	Não	$92,74 \pm 4,26$	$96,52 \pm 4,27$	$87,50 \pm 7,91$	$91,79 \pm 5,28$	$93,98 \pm 3,52$
100X	Não	Não	$92,89 \pm 5,06$	$100,00 \pm 0,00$	$82,50 \pm 12,75$	$89,81 \pm 6,69$	$94,50 \pm 3,72$
40X	Sim	Não	$96,89 \pm 2,54$	$98,33 \pm 3,33$	$95,00 \pm 6,12$	$96,67 \pm 4,08$	$97,39 \pm 2,13$
100X	Sim	Não	$94,89 \pm 3,16$	$100,00 \pm 0,00$	$87,50 \pm 7,91$	$92,27 \pm 4,55$	$95,92 \pm 2,44$
40X	Não	Sim	$91,68 \pm 6,37$	$92,88 \pm 6,77$	$90,00 \pm 9,35$	$93,16 \pm 6,71$	$92,87 \pm 5,53$
100X	Não	Sim	$93,89 \pm 3,72$	$98,33 \pm 3,33$	$87,50 \pm 7,91$	$92,14 \pm 4,56$	$95,06 \pm 2,97$
40X	Sim	Sim	$92,74 \pm 4,26$	$94,55 \pm 7,27$	$90,00 \pm 5,00$	$93,26 \pm 3,48$	$93,72 \pm 4,01$
100X	Sim	Sim	$96,84 \pm 4,21$	$98,18 \pm 3,64$	$95,00 \pm 6,12$	$96,52 \pm 4,27$	$97,31 \pm 3,62$

Fonte: elaborado pelo autor

De maneira complementar, para verificarmos a eficácia do uso de uma DenseNet como rede base, apresentamos nas Tabelas 5 e 6 os resultados dos experimentos com as arquiteturas Xception e Inception ResNet V2. O critério de ranqueamento escolhido para a decisão foi o F1-Score. Pode-se perceber, em relação aos resultados, que a DenseNet foi muito superior a essas outras arquiteturas de redes neurais convolucionais, justificando sua inclusão ao método proposto.

Tabela 5 – Comparativo de métricas em percentuais entre os diversos experimentos utilizando a arquitetura de CNN Xception.

Resultados dos Experimentos - Xception							
Amp.	Proc.	Aug.	Ac.	Sen.	Esp.	Prec.	F1-Score
40X	Não	Não	$93,74 \pm 6,15$	$98,18 \pm 3,64$	$87,50 \pm 13,69$	$92,36 \pm 7,86$	$94,98 \pm 4,83$
100X	Não	Não	$91,84 \pm 5,11$	$98,33 \pm 3,33$	$82,50 \pm 10,00$	$89,32 \pm 5,95$	$93,53 \pm 4,05$
40X	Sim	Não	$95,89 \pm 2,06$	$100,00 \pm 0,00$	$90,00 \pm 5,00$	$93,59 \pm 3,22$	$96,66 \pm 1,68$
100X	Sim	Não	$92,84 \pm 5,14$	$98,33 \pm 3,33$	$85,00 \pm 9,35$	$90,64 \pm 5,73$	$94,26 \pm 4,09$
40X	Não	Sim	$91,74 \pm 8,42$	$91,21 \pm 11,27$	$92,50 \pm 6,12$	$94,29 \pm 4,75$	$92,52 \pm 7,87$
100X	Não	Sim	$89,84 \pm 7,04$	$91,52 \pm 9,13$	$87,50 \pm 7,91$	$91,44 \pm 4,94$	$91,28 \pm 6,25$
40X	Sim	Sim	$93,79 \pm 6,07$	$93,03 \pm 8,55$	$95,00 \pm 6,12$	$96,33 \pm 4,52$	$94,46 \pm 5,51$
100X	Sim	Sim	$89,79 \pm 6,41$	$89,70 \pm 8,43$	$90,00 \pm 5,00$	$92,70 \pm 3,69$	$91,06 \pm 5,71$

Fonte: elaborado pelo autor

Tabela 6 – Comparativo de métricas em percentuais entre os diversos experimentos utilizando a arquitetura de CNN Inception ResNet V2.

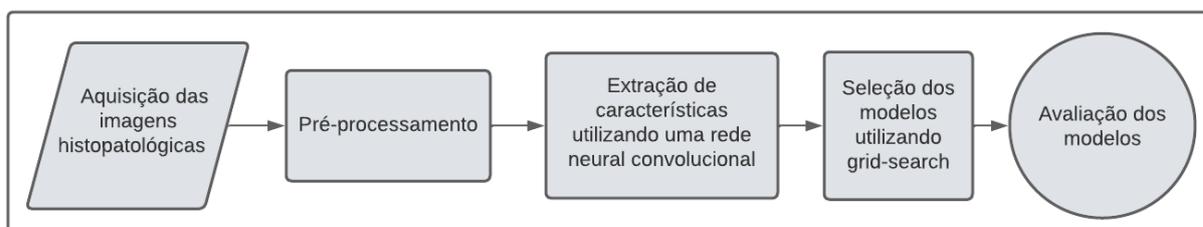
Resultados dos Experimentos - Inception ResNet V2							
Amp.	Proc.	Aug.	Ac.	Sen.	Esp.	Prec.	F1-Score
40X	Não	Não	91,84±6,79	93,18±9,75	90,00±5,00	92,98±3,60	92,88±6,23
100X	Não	Não	92,79±2,50	98,33±3,33	85,00±5,00	90,38±2,90	94,13±2,01
40X	Sim	Não	92,89±6,75	93,18±9,75	92,50±6,12	94,79±4,32	93,74±6,20
100X	Sim	Não	88,68±5,87	96,52±4,27	77,50±9,35	86,10±5,22	90,98±4,57
40X	Não	Sim	92,79±4,16	93,03±6,33	92,50±6,12	94,83±4,26	93,75±3,70
100X	Não	Sim	89,84±6,20	91,67±10,54	87,50±0,00	91,18±0,66	91,15±5,75
40X	Sim	Sim	94,84±3,33	94,70±4,34	95,00±6,12	96,64±4,13	95,56±2,89
100X	Sim	Sim	92,79±2,50	96,52±4,27	87,50±0,00	91,64±0,44	93,98±2,21

Fonte: elaborado pelo autor

5.2 Estudo de Ablação com a Técnica Deep Features

Com a finalidade de verificar os efeitos da transferência de aprendizado, a contribuição da especialização do treinamento dos pesos previamente obtidos e a escolha da rede classificadora no fim, realizamos um estudo no qual utilizou-se os pesos aprendidos originalmente para o *dataset* ImageNet e classificadores robustos apresentados na literatura. Todo o processo é otimizado para garantir o melhor resultado possível. A metodologia do estudo adicional é apresentada na Figura 19.

Figura 19 – Diagrama do método do estudo de ablação que utiliza as técnica de *Deep Features* em uma rede convolucional pré-treinada na ImageNet. As camadas convolucionais funcionam como um extrator de características.



Fonte: elaborado pelo autor

Utilizou-se as redes DenseNet, Inception ResNet V2 e a Xception pré-treinadas no *dataset* ImageNet, servindo como extratores de características das imagens histopatológicas melhoradas pelo algoritmo CLAHE. Neste caso, somente as camadas convolucionais previamente congeladas foram utilizadas. Os vetores de características resultantes serviram como entrada para um classificador nas etapas de treinamento e teste.

Para descobrir um bom modelo entre vários classificadores (*Decision Tree*, *Random Forest* e *K-Nearest Neighbors*), utilizou-se o algoritmo grid-search através de um processo de validação cruzada aninhada de 5x5 *folds* estratificada. O algoritmo seleciona por força

ht]

Tabela 7 – Tabela que descreve os parâmetros disponíveis no espaço de busca do algoritmo grid-search para cada classificador utilizado nos experimentos.

Classificador	Parâmetros
Decision Tree (DT)	Critério de qualidade da divisão: ["gini", "entropy"] Estratégia para a divisão: ["best", "random"]
Random Forest (RF)	Número de Árvores de Decisão: [200, 400, 600, 800] Critério de qualidade da divisão: ["gini", "entropy"]
K-Nearest Neighbors (KNN)	Número de k vizinhos: [10, 20, 40, 60, 80, 100] Distância como critério de peso: ["uniform", "distance"] Métrica utilizada: ["euclidean", "manhattan"]

Fonte: elaborado pelo autor

bruta os melhores parâmetros para cada classificador em cada rodada da validação cruzada. Tais parâmetros para cada classificador estão descritos na Tabela 7.

Para subsidiar estes experimentos, utilizou-se a biblioteca scikit-learn, que possui as implementações necessárias para o grid-search, os classificadores e para a geração de *folds*. Cada experimento teve seus resultados calculados conforme as métricas utilizadas para avaliação do método proposto, sendo agrupados por ampliação (amp.) e classificador (class.).

Os resultados relativos à rede DenseNet (Tabela 8) foram bastante promissores, em especial usando o classificador KNN para ampliação 40X, que teve seus indicadores ligeiramente superiores aos demais. Essa combinação obteve F1-Score de $95,47 \pm 5,23$. Para a ampliação 100X, o modelo treinado pelo classificador RF obteve um F1-Score de $95,71 \pm 2,89$.

Tabela 8 – Resultados da técnica *Deep Features* com a DenseNet.

Resultados dos Experimentos - DenseNet						
Amp.	Class.	Ac.	Sen.	Esp.	Prec.	F1-Score
40X	DT	87,68±3,99	91,36±9,14	82,50±15,00	89,59±8,63	89,69±3,42
40X	RF	93,84±1,92	98,33±3,33	87,50±7,91	92,14±4,56	94,98±1,34
40X	KNN	94,84±5,77	94,70±7,20	95,00±6,12	96,46±4,39	95,47±5,23
100X	DT	78,47±5,48	82,73±8,87	72,50±9,35	81,44±4,89	81,75±5,10
100X	RF	94,84±3,33	98,18±3,64	90,00±5,00	93,44±3,32	95,71±2,89
100X	KNN	91,84±8,27	94,70±7,20	87,50±13,69	92,18±7,90	93,27±6,78

Fonte: elaborado pelo autor

Com a Inception ResNet V2 (Tabela 9), o modelo selecionado treinado no classificador RF para ampliação 40X teve seus indicadores ligeiramente superiores aos demais, obtendo um F1-Score de $84,50 \pm 8,14$. Para a ampliação 100X, o modelo treinado pelo classificador RF obteve um F1-Score de $85,10 \pm 2,70$.

Tabela 9 – Resultados da técnica *Deep Features* com a Inception ResNet.

Resultados dos Experimentos - Inception ResNet V2						
Amp.	Class.	Ac.	Sen.	Esp.	Prec.	F1-Score
40X	DT	67,00±5,34	65,45±8,91	68,61±12,96	74,43±7,42	69,05±5,03
40X	RF	81,37±10,89	89,09±13,36	70,56±25,87	83,14±12,84	84,50±8,14
40X	KNN	78,32±10,74	80,00±14,55	75,28±22,52	84,19±12,81	80,44±9,30
100X	DT	70,11±9,68	69,09±4,45	71,39±18,16	77,77±13,43	72,80±7,64
100X	RF	82,47±2,49	89,09±8,91	73,33±12,74	82,44±5,04	85,10±2,70
100X	KNN	78,32±9,06	89,09±13,36	64,17±10,77	76,54±6,17	81,98±8,53

Fonte: elaborado pelo autor

Os resultados obtidos com a Xception (Tabela 10) foram inferiores em relação aos experimentos com a DenseNet e superiores aos da Inception ResNet V2. Mas no grupo de experimentos dessa rede convolucional, o modelo selecionado treinado no classificador RF para ampliação 40X teve seus indicadores ligeiramente superiores aos demais, obtendo um F1-Score de $88,24 \pm 6,32$. Para a ampliação 100X, o modelo treinado pelo classificador RF obteve um F1-Score de $88,83 \pm 4,34$.

Tabela 10 – Resultados da técnica *Deep Features* com a Xception.

Resultados dos Experimentos - Xception						
Amp.	Class.	Ac.	Sen.	Esp.	Prec.	F1-Score
40X	DT	76,21±15,88	78,18±16,86	73,06±32,08	83,47±16,25	78,84±12,71
40X	RF	85,47±9,16	92,73±6,80	75,56±20,95	85,26±11,03	88,24±6,32
40X	KNN	79,37±5,79	74,55±8,91	85,56±14,23	88,86±9,92	80,31±5,49
100X	DT	76,32±3,85	85,45±4,45	64,44±6,37	75,98±4,06	80,35±3,25
100X	RF	86,58±5,38	94,55±10,91	75,56±17,71	85,64±9,43	88,83±4,34
100X	KNN	85,63±8,78	89,09±6,80	81,11±16,98	87,32±11,14	87,75±7,09

Fonte: elaborado pelo autor

Os resultados mostraram a superioridade da DenseNet ao realizar as extrações de características e os melhores parâmetros encontrados pelo algoritmo grid-search estão listados na Tabela 11. Mas apesar destes resultados significativos, o método proposto se saiu melhor em relação ao método baseado em *Deep Features*, conforme o comparativo apresentado na Tabela 12.

Com base nos resultados alcançados, a DenseNet do método proposto aprendeu as características mais específicas das imagens histopatológicas na etapa de *fine-tuning* trazendo um melhor resultado. Já a DenseNet pré-treinada na ImageNet, utilizada na técnica de *Deep Features*, não conseguiu fornecer as características mais específicas que trariam uma melhor distinção entre classes aos classificadores.

Tabela 11 – Melhores parâmetros encontrados pelo algoritmo grid-search para os melhores resultados dos experimentos realizados.

Experimento	Parâmetros
KNN + DenseNet + imagens da ampliação 40X	Fold 1 -> Métrica: "manhattan", k: 40, pesos: "distance"
	Fold 2 -> Métrica: "manhattan", k: 20, pesos: "uniform"
	Fold 3 -> Métrica: "manhattan", k: 40, pesos: "uniform"
	Fold 4 -> Métrica: "manhattan", k: 10, pesos: "uniform"
	Fold 5 -> Métrica: "euclidean", k: 10, pesos: "distance"
RF + DenseNet + imagens da ampliação 100X	Fold 1 -> Número de árvores: 200, Critério de divisão: "gini"
	Fold 2 -> Número de árvores: 400, Critério de divisão: "gini"
	Fold 3 -> Número de árvores: 200, Critério de divisão: "entropy"
	Fold 4 -> Número de árvores: 400, Critério de divisão: "gini"
	Fold 5 -> Número de árvores: 400, Critério de divisão: "gini"

Fonte: elaborado pelo autor

Tabela 12 – Comparação entre o método proposto (linha 1 e 2) e o método baseado em *Deep Features* desenvolvido pelo mesmo autor (linha 3 e 4).

Método	Ampliação	Extração	Classificador	F1-Score(%)
Proposto	40X	DenseNet	DenseNet	97,39±2,13
Proposto	100X	DenseNet	DenseNet	97,31±3,62
Deep Features	40X	DenseNet	KNN	95,47±5,23
Deep Features	100X	DenseNet	RF	95,71±2,89

Fonte: elaborado pelo autor

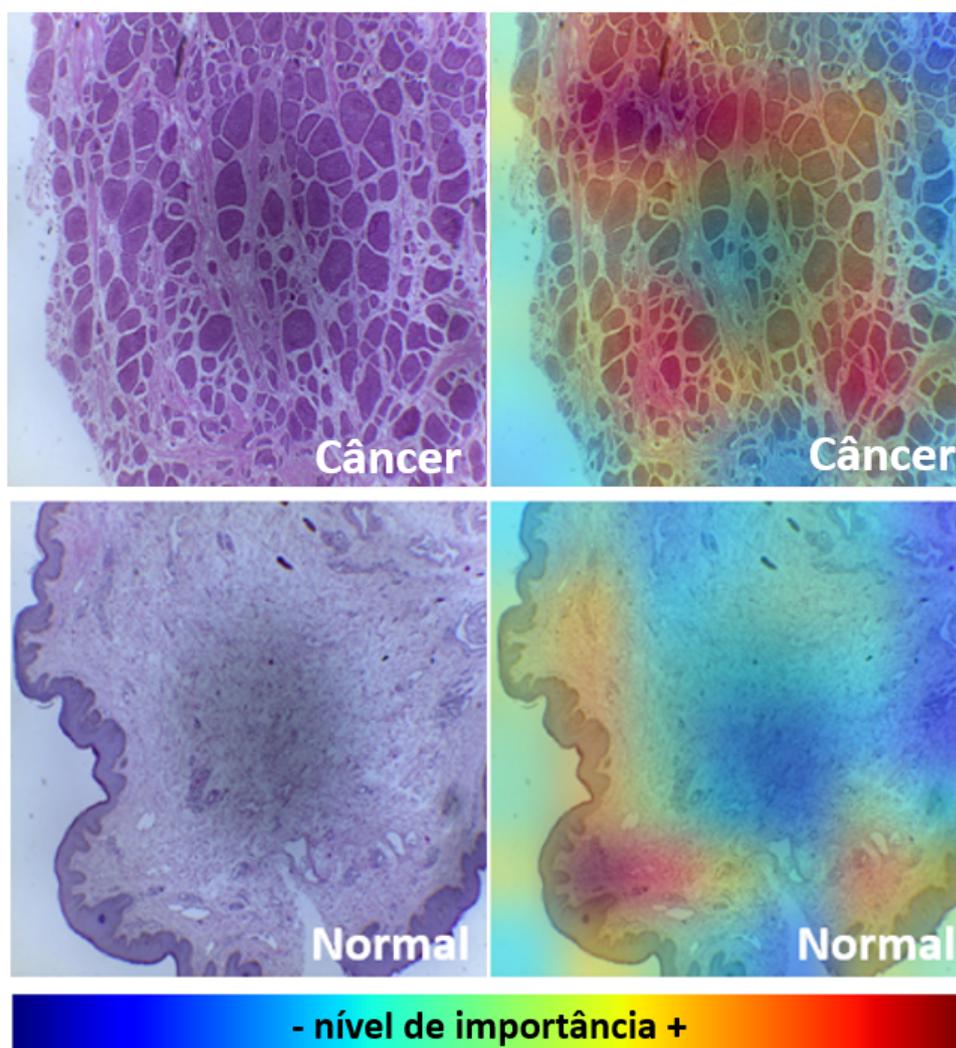
5.3 Explicabilidade do Modelo Gerado pelo Método Proposto

Para verificar a explicabilidade do modelo treinado, algumas amostras da partição de teste foram selecionadas, tanto para o caso positivo (câncer) quanto para o negativo (normal). Estes serviram de entrada para a rede DenseNet e os dados de ativação da última camada convolucional serviram para gerar o mapa de calor sobreposto a imagem original conforme mencionado no Capítulo 4.

Exemplos destes experimentos são apresentados na Figura 20 para as imagens de ampliação 40X e na Figura 21 para as de ampliação 100X.

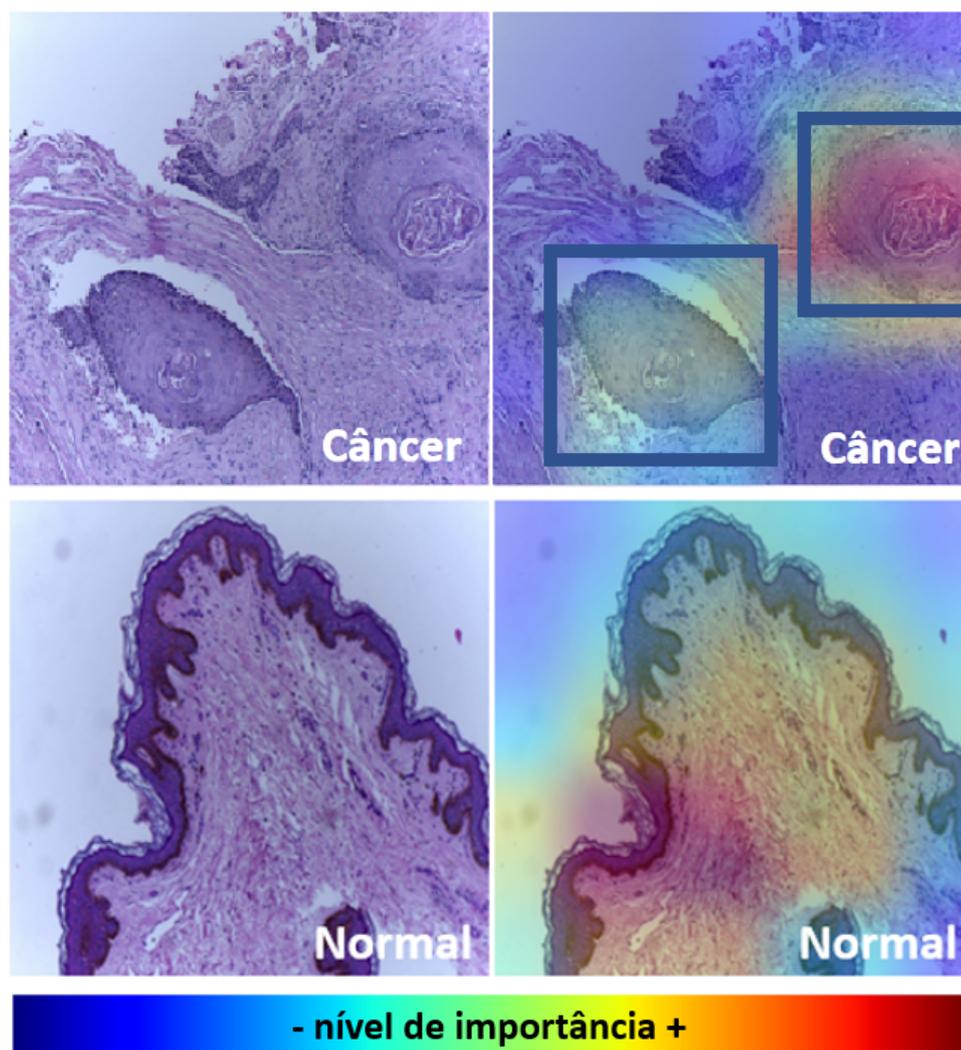
Com base nas amostras utilizadas para este experimento e com o apoio de um profissional patologista, chegou-se a conclusão que a rede neural considerou como mais importantes, na decisão de classificação, as regiões de borda dos tecidos para os casos considerados normais e para os constatados com câncer, as regiões com pérolas de queratina e agrupamentos celulares atípicos, isso pode ser notado nas marcações retangulares realizadas na imagem com câncer da Figura 21.

Figura 20 – Resultado do algoritmo Grad-CAM ao sobrepor os mapas de calor com as imagens de ampliação 40X avaliadas.



Fonte: elaborado pelo autor

Figura 21 – Resultado do algoritmo Grad-CAM ao sobrepor os mapas de calor com as imagens de ampliação 100X avaliadas.



Fonte: elaborado pelo autor

6 Conclusão

Neste trabalho foi apresentado um método baseado em redes neurais convolucionais e transferência de aprendizado para o problema de classificação binária de imagens histopatológicas de câncer de pênis. Apesar da pequena quantidade de amostras, que pode ser um fator limitante para o treinamento de um algoritmo de Aprendizado Profundo, o método alcançou resultados promissores com um F1-Score de até $97,39 \pm 2,13$ em uma base de imagens utilizada pela primeira vez em experimentos para automatização de diagnósticos médicos baseados em imagens. Foram produzidos experimentos, comparativos e discussões de resultados a respeito da aplicação de técnicas bastante populares em métodos baseados em Aprendizado Profundo como transferência de aprendizado, *fine-tuning*, *data augmentation* e *Deep Features*, sendo este último aplicado em um estudo de ablação. Além disso, a aplicação do algoritmo Grad-CAM para explicabilidade dos modelos gerados foi avaliada, produzindo imagens do tipo mapa de calor capazes de explicar quais regiões nas imagens histopatológicas foram mais interessantes para os modelos na tarefa de classificação.

Como trabalhos futuros, pretende-se utilizar técnicas que possam aproveitar melhor as características das imagens de alta resolução, como formas de extração por *patches* e a utilização de aprendizado fracamente supervisionado através de técnicas de *multiple instance learning*.

Existem ainda continuações no trabalho, prevista pelo grupo de profissionais da saúde do Projeto Câncer de Pênis da Amazônia Legal, que tem como objetivo ampliar o *dataset* com novas imagens, estudar a viabilidade de torná-lo público e adicionar informações que possibilitarão evoluir o método apresentado com o objetivo de classificar as imagens com câncer pela presença de HPV e grau histológico. Isto tornará o método mais robusto de forma a contribuir mais ainda com a saúde da população masculina. O autor almeja que futuramente estes estudos possibilitem desenvolver uma ferramenta que ajude os médicos a analisarem de forma mais precoce e assertiva o diagnóstico da doença, aproveitando todos os benefícios da patologia digital que supera barreiras de distância e supre a falta de profissionais em locais mais carentes, onde grupos de risco possuem menos acesso a serviços de saúde especializados.

Esta dissertação gerou uma publicação de seus resultados científicos, conforme informado na Tabela 13. O método proposto foi publicada no VISAPP (2022), evento A3 na Ciência da Computação.

Tabela 13 – Artigo publicado em evento científico em relação ao tema diagnóstico de câncer de pênis utilizando imagens histopatológicas.

Título do artigo	Congresso	Qualis
Classification of Histopathological Images of Penile Cancer using DenseNet and Transfer Learning (LAUANDE et al., 2022)	VISAPP 2022	A3

Fonte: elaborado pelo autor

Referências

- ACS, A. C. S. *Tests for Penile Cancer - American Cancer Society*. 2021. Acessado em 10/06/2021. Disponível em: <<https://www.cancer.org/cancer/penile-cancer>>. Citado 3 vezes nas páginas 15, 22 e 23.
- AL-JANABI, S.; HUISMAN, A.; DIEST, P. J. V. Digital pathology: current status and future perspectives. *Histopathology*, v. 61, n. 1, p. 1–9, 2012. Acessado em 12/03/2022. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2559.2011.03814.x>>. Citado na página 26.
- ANGHEL, A.; STANISAVLJEVIC, M.; ANDANI, S.; PAPANDREOU, N.; RÜSCHOFF, J. H.; WILD, P.; GABRANI, M.; POZIDIS, H. A high-performance system for robust stain normalization of whole-slide images in histopathology. *Frontiers in Medicine*, v. 6, 2019. ISSN 2296-858X. Acessado em 17/04/2022. Disponível em: <<https://www.frontiersin.org/article/10.3389/fmed.2019.00193>>. Citado 2 vezes nas páginas 19 e 21.
- ARAÚJO, T.; ARESTA, G.; CASTRO, E.; ROUCO, J.; AGUIAR, P.; ELOY, C.; POLONIA, A.; CAMPILHO, A. Classification of breast cancer histology images using convolutional neural networks. *PLoS one*, Public Library of Science San Francisco, CA USA, v. 12, n. 6, p. e0177544, 2017. Citado na página 20.
- BARROS, M. M. C. B. d. M. Érika Neves de. Câncer de pênis: perfil sócio-demográfico e respostas emocionais à penectomia em pacientes atendidos no serviço de psicologia do hospital de câncer de pernambuco. *Revista da SBPH*, sciELOpepsic, v. 12, p. 99 – 111, 06 2009. ISSN 1516-0858. Acessado em 12/08/2021. Disponível em: <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1516-08582009000100008&nrm=iso>. Citado na página 22.
- BEHRENS, H.-M.; WARNEKE, V. S.; BÖGER, C.; GARBRECHT, N.; JÜTTNER, E.; KLAPPER, W.; MATHIAK, M.; OSCHLIES, I.; RUDOLPH, U.; STUHLMANN-LAEISZ, C. et al. Reproducibility of her2/neu scoring in gastric cancer and assessment of the 10% cut-off rule. *Cancer medicine*, Wiley Online Library, v. 4, n. 2, p. 235–244, 2015. Citado na página 18.
- BOLHASANI, H.; AMJADI, E.; TABATABAEIAN, M.; JASSBI, S. J. A histopathological image dataset for grading breast invasive ductal carcinomas. *Informatics in Medicine Unlocked*, v. 19, p. 100341, 2020. ISSN 2352-9148. Acessado em 17/04/2022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2352914820300757>>. Citado 2 vezes nas páginas 20 e 21.
- BORG, M.; JABANGWE, R.; ÅBERG, S.; EKBLÖM, A.; HEDLUND, L.; LIDFELDT, A. Test automation with grad-cam heatmaps - a future pipe segment in ml ops for vision ai? In: *2021 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*. [S.l.: s.n.], 2021. p. 175–181. Citado na página 46.
- BOUMARAF, S.; LIU, X.; ZHENG, Z.; MA, X.; FERKOUS, C. A new transfer learning based approach to magnification dependent and independent classification of

- breast cancer in histopathological images. *Biomedical Signal Processing and Control*, v. 63, p. 102192, 2021. ISSN 1746-8094. Acessado em 27/09/2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S174680942030330X>>. Citado 2 vezes nas páginas 20 e 21.
- CHAN, L.; MARINKOVICH, M. *New Technologies in Dermatological Science and Practice: Intelligence, Regeneration, Speed, and Precision*. [S.l.]: CRC Press, 2021. ISBN 9781000460728. Citado 2 vezes nas páginas 26 e 27.
- CHOLLET, F. *Xception: Deep Learning with Depthwise Separable Convolutions*. arXiv, 2016. Acessado em 27/03/2022. Disponível em: <<https://arxiv.org/abs/1610.02357>>. Citado na página 43.
- CHOLLET, F. *Deep Learning with Python, Second Edition*. [S.l.]: Manning, 2021. 7 p. ISBN 9781617296864. Citado 2 vezes nas páginas 27 e 31.
- CHOLLET, F. et al. *Keras*. 2015. Acessado em 02/05/2021. Disponível em: <<https://keras.io>>. Citado na página 47.
- CHOUDHARY, T.; MISHRA, V.; GOSWAMI, A.; SARANGAPANI, J. A transfer learning with structured filter pruning approach for improved breast cancer classification on point-of-care devices. *Computers in Biology and Medicine*, v. 134, p. 104432, 2021. ISSN 0010-4825. Acessado em 09/07/2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010482521002262>>. Citado 2 vezes nas páginas 20 e 21.
- CIOMPI, F.; GEESINK, O.; BEJNORDI, B. E.; SOUZA, G. S. de; BAIDOSHVILI, A.; LITJENS, G.; GINNEKEN, B. van; NAGTEGAAL, I.; LAAK, J. van der. The importance of stain normalization in colorectal tissue classification with convolutional networks. In: *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. [S.l.: s.n.], 2017. p. 160–163. Citado 2 vezes nas páginas 19 e 21.
- COELHO, R.; PINHO, J.; MORENO, J.; GARBIS, D.; NASCIMENTO, A.; LARGES, J.; CALIXTO, J.; RAMALHO, L.; SILVA, A.; NOGUEIRA, L.; FEITOZA, L.; BARROS-SILVA, G. Penile cancer in maranhão, northeast brazil: The highest incidence globally? *BMC Urology*, v. 18, 05 2018. Citado na página 15.
- COSTA, S.; RODRIGUES, R.; BARBOSA, L.; SILVA, J.; BRANDÃO, J. O. d. C.; MEDEIROS, C. S. Q. d. Câncer de pênis: epidemiologia e estratégias de prevenção. *Caderno de Graduação - Ciências Biológicas e da Saúde - UNIT - PERNAMBUCO*, v. 1, n. 2, p. 23–33, nov. 2013. Acessado em 27/03/2022. Disponível em: <<https://periodicos.set.edu.br/facipesaude/article/view/1197>>. Citado na página 23.
- CRUZ-ROA, A.; BASAVANHALLY, A.; GONZÁLEZ, F.; GILMORE, H.; FELDMAN, M.; GANESAN, S.; SHIH, N.; TOMASZEWSKI, J.; MADABHUSHI, A. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In: GURCAN, M. N.; MADABHUSHI, A. (Ed.). *Medical Imaging 2014: Digital Pathology*. SPIE, 2014. v. 9041, p. 1 – 15. Acessado em 08/06/2021. Disponível em: <<https://doi.org/10.1117/12.2043872>>. Citado na página 16.
- CUBUK, E. D.; ZOPH, B.; SHLENS, J.; LE, Q. Randaugment: Practical automated data augmentation with a reduced search space. In: LAROCHELLE, H.; RANZATO,

- M.; HADSELL, R.; BALCAN, M. F.; LIN, H. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2020. v. 33, p. 18613–18624. Acessado em 27/03/2022. Disponível em: <<https://proceedings.neurips.cc/paper/2020/file/d85b63ef0ccb114d0a3bb7b7d808028f-Paper.pdf>>. Citado na página 20.
- DIGABEL, H.; LANTUÉJOUL, C. Iterative algorithms. In: VERLAG, R. (Ed.). *Actes du Second Symposium Europeen d'Analyse Quantitative des Microstructures en Sciences des Materiaux, Biologie et Medecine*. [S.l.: s.n.], 1978. p. 85–99. Citado na página 19.
- DOUGLAWI, A.; MASTERSON, T. Updates on the epidemiology and risk factors for penile cancer. *Translational Andrology and Urology*, v. 6, p. 785–790, 10 2017. Citado na página 15.
- ERDMANN, M.; GLOMBITZA, J.; KASIECZKA, G.; KLEMRADT, U. *Deep Learning For Physics Research*. [S.l.]: World Scientific Publishing Company, 2021. 76, 77, 233, 234 p. ISBN 9789811237478. Citado na página 33.
- ESTEVA, A.; CHOU, K.; YEUNG, S.; NAIK, N.; MADANI, A.; MOTTAGHI, A.; LIU, Y.; TOPOL, E.; DEAN, J.; SOCHER, R. Deep learning-enabled medical computer vision. *npj Digital Medicine*, v. 4, p. 5, 12 2021. Citado na página 27.
- FARAHANI, N.; PANTANOWITZ, L. Overview of telepathology. *Surgical Pathology Clinics*, v. 8, n. 2, p. 223–231, 2015. ISSN 1875-9181. Pathology Informatics. Acessado em 12/06/2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1875918115000367>>. Citado na página 26.
- FARYNA, K.; LAAK, J. van der; LITJENS, G. Tailoring automated data augmentation to he-stained histopathology. In: HEINRICH, M.; DOU, Q.; BRUIJNE, M. de; LELLMANN, J.; SCHLÄFER, A.; ERNST, F. (Ed.). *Proceedings of the Fourth Conference on Medical Imaging with Deep Learning*. PMLR, 2021. (Proceedings of Machine Learning Research, v. 143), p. 168–178. Acessado em 02/04/2022. Disponível em: <<https://proceedings.mlr.press/v143/faryna21a.html>>. Citado 2 vezes nas páginas 20 e 21.
- FILIPCZUK, P.; FEVENS, T.; KRZYŻAK, A.; MONCZAK, R. Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. *IEEE Transactions on Medical Imaging*, v. 32, n. 12, p. 2169–2178, 2013. Citado 2 vezes nas páginas 18 e 21.
- GÉRON, A. *Mãos A Obra: Aprendizado De Máquina Com Scikit-Learn, Keras E TensorFlow: Conceitos, Ferramentas e Técnicas Para a Construção de Sistemas Inteligentes*. [S.l.]: Alta Books, 2021. ISBN 9786555208146. Citado na página 31.
- GRZYBOWSKI, A. *Artificial Intelligence in Ophthalmology*. [S.l.]: Springer International Publishing, 2021. 72 p. ISBN 9783030786014. Citado na página 32.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. *Deep Residual Learning for Image Recognition*. arXiv, 2015. Acessado em 27/03/2022. Disponível em: <<https://arxiv.org/abs/1512.03385>>. Citado na página 34.
- HEBB, D. *The Organization of Behavior: A Neuropsychological Theory*. [S.l.]: Wiley, 1949. (Neuropsychological Theory). ISBN 9780471367277. Citado na página 28.

HINTON, G. E.; SALAKHUTDINOV, R. R. Reducing the dimensionality of data with neural networks. *science*, American Association for the Advancement of Science, v. 313, n. 5786, p. 504–507, 2006. Citado na página 31.

HINTON, G. E.; SRIVASTAVA, N.; KRIZHEVSKY, A.; SUTSKEVER, I.; SALAKHUTDINOV, R. R. *Improving neural networks by preventing co-adaptation of feature detectors*. arXiv, 2012. Acessado em 19/02/2022. Disponível em: <<https://arxiv.org/abs/1207.0580>>. Citado na página 42.

HUANG, G.; LIU, Z.; MAATEN, L. V. D.; WEINBERGER, K. Q. Densely connected convolutional networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2017. p. 2261–2269. Citado 3 vezes nas páginas 8, 34 e 35.

HUSS, R.; GRUNKIN, M. *Artificial Intelligence Applications In Human Pathology*. [S.l.]: World Scientific Publishing Company, 2022. ISBN 9781800611405. Citado 2 vezes nas páginas 27 e 28.

IBGE. *Panorama - Estado do Maranhão - Instituto Brasileiro de Geografia e Estatística - IBGE*. 2022. Acessado em 12/03/2022. Disponível em: <<https://cidades.ibge.gov.br/brasil/ma/panorama>>. Citado na página 15.

INCA. *Types of Cancer | National Cancer Institute - José Alencar Gomes da Silva - INCA*. National Cancer Institute - José Alencar Gomes da Silva - INCA, 2021. Acessado em 10/06/2021. Disponível em: <<https://www.inca.gov.br/tipos-de-cancer/cancer-de-penis>>. Citado 2 vezes nas páginas 15 e 23.

KATHER, J. N.; WEIS, C.-A.; BIANCONI, F.; MELCHERS, S. M.; SCHAD, L. R.; GAISER, T.; MARX, A.; ZÖLLNER, F. G. Multi-class texture analysis in colorectal cancer histology. *Scientific reports*, Nature Publishing Group, v. 6, n. 1, p. 1–11, 2016. Citado na página 21.

KETKAR, N. *Deep Learning with Python: A Hands-on Introduction*. [S.l.]: Apress, 2017. ISBN 9781484227664. Citado 2 vezes nas páginas 28 e 31.

KIANI, A.; UYUMAZTURK, B.; RAJPURKAR, P.; WANG, A.; GAO, R.; JONES, E.; YU, Y.; LANGLOTZ, C.; BALL, R.; MONTINE, T.; MARTIN, B.; BERRY, G.; OZAWA, M.; HAZARD, F.; BROWN, R.; CHEN, S.; WOOD, M.; ALLARD, L.; YLAGAN, L.; SHEN, J. Impact of a deep learning assistant on the histopathologic classification of liver cancer. *npj Digital Medicine*, v. 3, 12 2020. Citado na página 16.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*. Red Hook, NY, USA: Curran Associates Inc., 2012. (NIPS'12), p. 1097–1105. Citado 2 vezes nas páginas 18 e 32.

KUMAR, A.; SHAIK, F. *Image Processing in Diabetic Related Causes*. [S.l.]: Springer Singapore, 2015. (SpringerBriefs in Applied Sciences and Technology). ISBN 9789812876249. Citado na página 41.

LAUANDE, M.; TELES, A.; Lima da Silva., L.; MATOS, C.; Braz Júnior, G.; Cardoso de Paiva, A.; Sousa de Almeida., J.; OLIVEIRA, R.; BRITO., H.; NASCIMENTO, A.; PESTANA., A.; SANTOS, A.; LOPES., F. Classification of histopathological images of

penile cancer using densenet and transfer learning. In: INSTICC. *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 4: VISAPP*, [S.l.]: SciTePress, 2022. p. 976–983. ISBN 978-989-758-555-5. Citado na página 56.

LECUN, Y.; BOSER, B.; DENKER, J. S.; HENDERSON, D.; HOWARD, R. E.; HUBBARD, W.; JACKEL, L. D. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, v. 1, n. 4, p. 541–551, 1989. Citado 2 vezes nas páginas 19 e 32.

LINKON, A. H. M.; LABIB, M. M.; HASAN, T.; HOSSAIN, M.; MARIUM-E-JANNAT. Deep learning in prostate cancer diagnosis and gleason grading in histopathology images: An extensive study. *Informatics in Medicine Unlocked*, v. 24, p. 100582, 2021. ISSN 2352-9148. Acessado em 19/06/2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2352914821000721>>. Citado na página 16.

MALAVADE, V. N.; MELINAMATH, B. C.; PARDESHI, S. A. Survey of mitosis detection techniques in breast cancer. In: *2018 3rd International Conference on Inventive Computation Technologies (ICICT)*. [S.l.: s.n.], 2018. p. 830–832. Citado na página 16.

MATIGNON, R. *Neural Network Modeling Using Sas Enterprise Miner*. [S.l.]: AuthorHouse, 2005. 101, 102 p. ISBN 9781418423414. Citado na página 29.

MCCULLOCH, W.; PITTS, W. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, v. 5, p. 127–147, 1943. Citado na página 28.

MELO, R. C. N.; RAAS, M. W. D.; PALAZZI, C.; NEVES, V. H.; MALTA, K. K.; SILVA, T. P. Whole slide imaging and its applications to histopathological studies of liver disorders. *Frontiers in Medicine*, v. 6, p. 310, 2020. ISSN 2296-858X. Acessado em 19/03/2022. Disponível em: <<https://www.frontiersin.org/article/10.3389/fmed.2019.00310>>. Citado na página 15.

MISHRA, A. *Machine Learning for iOS Developers*. [S.l.]: Wiley, 2020. 85 p. ISBN 9781119602910. Citado na página 32.

MORID, M. A.; BORJALI, A.; Del Fiol, G. A scoping review of transfer learning research on medical image analysis using imagenet. *Computers in Biology and Medicine*, v. 128, p. 104115, 2021. ISSN 0010-4825. Acessado em 19/06/2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010482520304467>>. Citado na página 37.

MOSCONI, A. M.; ROILA, F.; GATTA, G.; THEODORE, C. Cancer of the penis. *Critical Reviews in Oncology/Hematology*, v. 53, n. 2, p. 165–177, 2005. ISSN 1040-8428. Acessado em 19/06/2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1040842804001593>>. Citado na página 23.

NCI, N. C. I. *Definition of corpus spongiosum - NCI Dictionary of Cancer Terms - National Cancer Institute*. 2022. Acessado em 19/06/2021. Disponível em: <<https://www.cancer.gov/publications/dictionaries/cancer-terms/def/corpus-spongiosum>>. Citado na página 22.

- NETO, A. D. *Cytopathology technician: reference book 3: Histopathology techniques / Brasília; Ministry of Health of Brazil ; 2012. 93 p.* / MS. Ministry of Health of Brazil, 2012. 11–25 p. Acessado em 19/06/2021. Disponível em: <<https://pesquisa.bvsalud.org/bvsms/resource/pt/mis-37140>>. Citado na página 24.
- NIAZI, M. K. K.; PARWANI, A. V.; GURCAN, M. N. Digital pathology and artificial intelligence. *The Lancet Oncology*, v. 20, n. 5, p. e253–e261, 2019. ISSN 1470-2045. Acessado em 12/06/2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1470204519301548>>. Citado na página 27.
- NUNES, C. de S.; CINSA, L. A. Princípios do processamento histológico de rotina. *Revista Interdisciplinar de Estudos Experimentais-Animais e Humanos Interdisciplinary Journal of Experimental Studies*, v. 8, n. 1, 2016. Citado na página 25.
- PAL, K. K.; SUDEEP, K. S. Preprocessing for image classification by convolutional neural networks. In: *2016 IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)*. [S.l.: s.n.], 2016. p. 1778–1781. Citado na página 41.
- PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011. Citado na página 47.
- POMPEO, A.; KOFF, W.; DAMIÃO, R. et al. Carcinoma do pênis-parte ii. *Projeto Diretrizes–Associação Médica Brasileira e Conselho Federal de Medicina*, p. 1–9, 2006. Citado na página 23.
- RAKHLIN, A.; SHVETS, A.; IGLOVIKOV, V.; KALININ, A. A. Deep convolutional neural networks for breast cancer histology image analysis. In: CAMPILHO, A.; KARRAY, F.; ROMENY, B. ter H. (Ed.). *Image Analysis and Recognition*. Cham: Springer International Publishing, 2018. p. 737–744. Citado 2 vezes nas páginas 20 e 21.
- RASCHKA, S.; MIRJALILI, V. *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2, 3rd Edition*. [S.l.]: Packt Publishing, 2019. 383 p. ISBN 9781789958294. Citado na página 31.
- RODRIGUES, V. *como Máquinas aprendem: Volume II ...algoritmos clássicos de Machine Learning*. [S.l.]: Independently Published, 2021. 467 p. ISBN 9798720965839. Citado na página 30.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, v. 65 6, p. 386–408, 1958. Citado na página 28.
- ROY, S.; kumar Jain, A.; LAL, S.; KINI, J. A study about color normalization methods for histopathology images. *Micron*, v. 114, p. 42–61, 2018. ISSN 0968-4328. Acessado em 17/04/2022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0968432818300982>>. Citado na página 19.

- RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA, S.; HUANG, Z.; KARPATY, A.; KHOSLA, A.; BERNSTEIN, M. et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, Springer, v. 115, n. 3, p. 211–252, 2015. Citado na página 19.
- SAMAH, A. A.; FAUZI, M. F. A.; MANSOR, S. Classification of benign and malignant tumors in histopathology images. In: *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*. [S.l.: s.n.], 2017. p. 102–106. Citado 2 vezes nas páginas 18 e 21.
- SANTOS, I. V.; BOMFIM, A. M. A.; LESSA, V. J. C.; FERREIRA, A. M. V. Estratégias do enfermeiro na prevenção do câncer de pênis. *Revista Uniabeu*, v. 11, n. 29, p. 362–373, 2018. Citado na página 22.
- SARWINDA, D.; PARADISA, R. H.; BUSTAMAM, A.; ANGGIA, P. Deep learning in image classification using residual network (resnet) variants for detection of colorectal cancer. *Procedia Computer Science*, v. 179, p. 423–431, 2021. ISSN 1877-0509. 5th International Conference on Computer Science and Computational Intelligence 2020. Acessado em 12/06/2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050921000284>>. Citado 3 vezes nas páginas 16, 19 e 21.
- SELVARAJU, R. R.; COGSWELL, M.; DAS, A.; VEDANTAM, R.; PARIKH, D.; BATRA, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision*, Springer Science and Business Media LLC, v. 128, n. 2, p. 336–359, Oct 2019. ISSN 1573-1405. Acessado em 19/06/2021. Disponível em: <<http://dx.doi.org/10.1007/s11263-019-01228-7>>. Citado na página 46.
- SHARMA, H.; ZERBE, N.; KLEMPERT, I.; HELLWICH, O.; HUFNAGL, P. Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. *Computerized Medical Imaging and Graphics*, v. 61, p. 2–13, 2017. ISSN 0895-6111. Selected papers from the 13th European Congress on Digital Pathology. Acessado em 19/03/2022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0895611117300502>>. Citado 2 vezes nas páginas 18 e 21.
- SHEN, D.; WU, G.; SUK, H.-I. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, v. 19, n. 1, p. 221–248, 2017. PMID: 28301734. Acessado em 19/03/2022. Disponível em: <<https://doi.org/10.1146/annurev-bioeng-071516-044442>>. Citado na página 27.
- SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of Big Data*, Springer, v. 6, n. 1, p. 1–48, 2019. Citado 2 vezes nas páginas 36 e 37.
- SILVA, H.; CORREIA, L. *Manual Eletrônico sobre Artefatos Durante a Preparação de Lâminas Histológicas*. [S.l.]: Editora Appris, 2021. ISBN 9786525010038. Citado na página 24.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv 1409.1556*, 09 2014. Citado na página 19.

- SPANHOL, F. A.; OLIVEIRA, L. S.; CAVALIN, P. R.; PETITJEAN, C.; HEUTTE, L. Deep features for breast cancer histopathological image classification. In: *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. [S.l.: s.n.], 2017. p. 1868–1873. Citado 2 vezes nas páginas 19 e 21.
- SPANHOL, F. A.; OLIVEIRA, L. S.; PETITJEAN, C.; HEUTTE, L. Breast cancer histopathological image classification using convolutional neural networks. In: *2016 International Joint Conference on Neural Networks (IJCNN)*. [S.l.: s.n.], 2016. p. 2560–2567. Citado 2 vezes nas páginas 18 e 21.
- SPANHOL, F. A.; OLIVEIRA, L. S.; PETITJEAN, C.; HEUTTE, L. A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, v. 63, n. 7, p. 1455–1462, 2016. Citado 2 vezes nas páginas 18 e 21.
- SRINIDHI, C. L.; CIGA, O.; MARTEL, A. L. Deep neural network models for computational histopathology: A survey. *Medical Image Analysis*, v. 67, p. 101813, 2021. ISSN 1361-8415. Acessado em 12/06/2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1361841520301778>>. Citado na página 16.
- SZEGEDY, C.; IOFFE, S.; VANHOUCKE, V.; ALEMI, A. *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning*. arXiv, 2016. Acessado em 22/03/2022. Disponível em: <<https://arxiv.org/abs/1602.07261>>. Citado na página 43.
- TAJBAKHS, N.; SHIN, J. Y.; GURUDU, S. R.; HURST, R. T.; KENDALL, C. B.; GOTWAY, M. B.; LIANG, J. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging*, v. 35, n. 5, p. 1299–1312, May 2016. ISSN 1558-254X. Citado na página 37.
- TAN, H. H.; LIM, K. H. Vanishing gradient mitigation with deep learning neural network optimization. In: *2019 7th International Conference on Smart Computing Communications (ICSCC)*. [S.l.: s.n.], 2019. p. 1–4. Citado na página 34.
- TELLEZ, D.; LITJENS, G.; BÁNDI, P.; BULTEN, W.; BOKHORST, J.-M.; CIOMPI, F.; van der Laak, J. Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology. *Medical Image Analysis*, v. 58, p. 101544, 2019. ISSN 1361-8415. Acessado em 09/07/2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1361841519300799>>. Citado 2 vezes nas páginas 20 e 21.
- THOMAS, A.; NECCHI, A.; MUNEER, A.; TOBIAS-MACHADO, M.; TRAN, A. T. H.; ROMPUY, A.-S. V.; SPIESS, P. E.; ALBERSEN, M. Penile cancer. *Nature Reviews Disease Primers*, Nature Publishing Group, v. 7, n. 1, p. 1–24, 2021. Citado na página 15.
- VIDYARTHI, A.; SHAD, J.; SHARMA, S.; AGARWAL, P. Classification of breast microscopic imaging using hybrid clahe-cnn deep architecture. In: *2019 Twelfth International Conference on Contemporary Computing (IC3)*. [S.l.: s.n.], 2019. p. 1–5. Citado 3 vezes nas páginas 16, 19 e 21.
- VIEIRA, C.; FEITOZA, L.; PINHO, J.; TEIXEIRA-JÚNIOR, A.; LAGES, J.; CALIXTO, J.; COELHO, R.; NOGUEIRA, L.; CUNHA, I.; SOARES, F.; BARROS-SILVA, G. Profile of patients with penile cancer in the region with the highest worldwide incidence. *Scientific Reports*, v. 10, 02 2020. Citado na página 15.

- VILLÁN, A. *Mastering OpenCV 4 with Python: A practical guide covering topics from image processing, augmented reality to deep learning with OpenCV 4 and Python 3.7*. [S.l.]: Packt Publishing, 2019. 185 p. ISBN 9781789349757. Citado na página 42.
- WEI, J.; TAFE, L.; LINNIK, Y.; VAICKUS, L.; TOMITA, N.; HASSANPOUR, S. Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. *Scientific Reports*, v. 9, 03 2019. Citado na página 16.
- WIND, M. M.; FERNANDES, L. M. S.; PINHEIRO, D. H. P.; FERREIRA, V. R.; GABRIEL, A. C. G.; SILVA, C. T. X. Perfil epidemiológico do câncer de pênis e suas consequências psíquicas. *CIPEEX*, v. 2, p. 932–936, 2018. Citado na página 23.
- YANG, Y.; YUHUA, P.; ZHAOGUANG, L. A fast algorithm for ycbcr to rgb conversion. *IEEE Transactions on Consumer Electronics*, v. 53, n. 4, p. 1490–1493, 2007. Citado na página 42.
- ZHANG, A.; LIPTON, Z. C.; LI, M.; SMOLA, A. J. Dive into deep learning. *arXiv preprint arXiv:2106.11342*, 2021. Citado 2 vezes nas páginas 32 e 33.
- ZHOU, S. K.; GREENSPAN, H.; DAVATZIKOS, C.; DUNCAN, J. S.; GINNEKEN, B. V.; MADABHUSHI, A.; PRINCE, J. L.; RUECKERT, D.; SUMMERS, R. M. A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. *Proceedings of the IEEE*, v. 109, n. 5, p. 820–838, 2021. Citado na página 28.
- ZHOU, X.; LI, C.; RAHAMAN, M. M.; YAO, Y.; AI, S.; SUN, C.; WANG, Q.; ZHANG, Y.; LI, M.; LI, X.; JIANG, T.; XUE, D.; QI, S.; TENG, Y. A comprehensive review for breast histopathology image analysis using classical and deep neural networks. *IEEE Access*, v. 8, p. 90931–90956, 2020. Citado na página 16.
- ZUIDERVELD, K. J. Contrast limited adaptive histogram equalization. In: *Graphics Gems*. [S.l.: s.n.], 1994. Citado na página 41.