

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Higor Emanuel Souza Silva

**Classificação de Lesões em Imagens
Histológicas de Mama usando Wavelet e
ResNet-50**

Uberlândia, Brasil

2021

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Higor Emanuel Souza Silva

**Classificação de Lesões em Imagens Histológicas de
Mama usando Wavelet e ResNet-50**

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, Minas Gerais, como requisito exigido parcial à obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Marcelo Zanchetta do Nascimento

Universidade Federal de Uberlândia – UFU

Faculdade de Computação

Bacharelado em Ciência da Computação

Uberlândia, Brasil

2021

Resumo

O câncer de mama é um dos principais problemas de saúde pública no Brasil e no mundo, com uma taxa de mortalidade de 16%. Essa doença pode ser diagnosticada em suas fases iniciais, aumentando a possibilidade de tratamentos menos agressivos e com maior chance de sucesso. Existem diversas técnicas de diagnóstico como a mamografia, os raios-X e a análise de tecido histológico no microscópio. O diagnóstico é feito por especialistas, em processos que demandam muita atenção e que estão suscetíveis a erros humanos. Uma das maneiras de auxiliar estes médicos é a utilização de sistemas computacionais de auxílio ao diagnóstico, que são algoritmos computacionais especializados em classificar diferentes tipos de lesões. Neste trabalho, foram utilizadas imagens histológicas digitalizadas de lesões de mama da base de dados BreakHis. Dessas imagens, foram extraídos descritores *handcrafted* utilizando a transformada wavelet Daubechies 8 e *non-handcrafted* utilizando a ResNet-50. Foi analisada a combinação destes dois tipos de descritores, com os extraídos da camada de *flatten* da ResNet-50, combinados com os extraídos da transformada wavelet: detalhe, aproximação e informação de energia. Então, foram avaliadas as métricas acurácia e AUC da associação dos descritores utilizando os classificadores RF e SVM. A associação dos descritores não melhorou os resultados em relação ao *baseline*, e não houve diferença significativa em utilizar os descritores *handcrafted*.

Palavras-chave: Transformada *Wavelet*, ResNet-50, CAD, Classificação, Histologia.

Lista de ilustrações

Figura 1 – Anatomia da mama. Fonte: (CANCER, 2021)	11
Figura 2 – Exemplos de imagens com lesões de mama benignas e malignas. Fonte: (SPANHOL et al., 2016)	11
Figura 3 – Etapas empregadas em um sistema CAD - Fonte: (KAUSHAL et al., 2019)	13
Figura 4 – Exemplos de texturas em uma lesão maligna e uma ROI benigna. Fonte: (SPANHOL et al., 2016)	15
Figura 5 – Aplicação de uma Transformada Wavelet discreta em uma imagem histológica de mama, utilizando a função <i>wavelet</i> -mãe Daubechies 4. Fonte: (MOREIRA JUNIOR; NASCIMENTO, 2018)	17
Figura 6 – Representação de neurônios reais (esquerda) e artificiais (direita). Fonte: (ZUIN; VELOSO, 2017)	19
Figura 7 – Exemplo do modelo proposto por LeCun para reconhecimento de dígitos numéricos. Fonte: (AL, 1989)	20
Figura 8 – Exemplos de imagens benignas e malignas com múltiplos fatores de magnificação. Fonte: (SPANHOL et al., 2016)	24
Figura 9 – O modelo proposto combina os descritores <i>handcrafted</i> extraídos pela <i>wavelet</i> , com os descritores <i>non-handcrafted</i> extraídos pela ResNet-50.	25
Figura 10 – Representação das combinações de <i>features</i> produzidas no processo de classificação.	26
Figura 11 – A imagem mostra alguns exemplos de arquiteturas profundas, como a VGG (SIMONYAN; ZISSERMAN, 2014), e outras redes com e sem utilização de resíduo.	28
Figura 12 – Avaliação do processo de treinamento e teste na arquitetura ResNet-50 com diferentes níveis de magnificação das imagens histológicas.	32
Figura 13 – ECDF para as características extraídas do modelo ResNet-50.	32
Figura 14 – ECDF para as características extraída do modelo ResNet-50 combinadas com as sub-bandas de detalhe da transformada wavelet.	33
Figura 15 – ECDF para as características extraídas do modelo ResNet-50 combinadas com as sub-bandas de aproximação da transformada wavelet.	34
Figura 16 – ECDF para as características extraída do modelo ResNet50 combinadas com a energia extraída das sub-bandas de detalhe da aplicação da Transformada Wavelet Daubechies 8.	35
Figura 17 – ECDF para as características extraída do modelo ResNet-50 combinadas com a energia extraída da sub-banda de aproximação da aplicação da transformada wavelet Daubechies 8.	36

Lista de abreviaturas e siglas

AI	Artificial Intelligence
AUC	Area Under the Curve
CAD	Computer Aided Diagnosis
CNN	Convolutional Neural Network
CSDCNN	Class Structure-based Deep Convolutional Neural Networks
ECDF	Empirical Cumulative Distribution Function
INCA	Instituto Nacional do Câncer
ML	Machine Learning
PDI	Processamento Digital de Imagens
ReLU	Rectified Linear Unit
ResNet	Residual Networks
RF	Random Forest
RNA	Rede Neural Artificial
ROI	Region of Interest
SVM	Support Vector Machines
TF	Transformada de Fourier
TPJF	Transformada por Janela de Fourier
TW	Transformada Wavelet
VLAD	Vector of Locally Aggregated Descriptors

Sumário

1	INTRODUÇÃO	7
1.1	Objetivos	9
1.1.1	Geral	9
1.1.2	Específicos	9
1.2	Organização desse Documento	9
2	CONTEXTUALIZAÇÃO	10
2.1	Câncer de mama	10
2.2	Sistema de Apoio ao Diagnóstico	12
2.3	Processamento Digital de Imagens	12
2.3.1	Aquisição da Imagem	12
2.3.2	Pré-Processamento	14
2.3.3	Segmentação	14
2.3.4	Representação e Descrição	14
2.3.5	Transformada Wavelet	16
2.4	Redes Neurais Convolucionais	18
2.5	Trabalhos Correlatos	21
2.6	Considerações Finais	22
3	MATERIAIS E MÉTODOS	23
3.1	Base de Dados	23
3.2	Modelo Proposto	23
3.3	Descritores	25
3.4	ResNet-50	26
3.5	Transformada <i>Wavelet</i>	27
3.6	Classificação	29
3.6.1	Random Forest	29
3.6.2	Support Vector Machine	29
3.7	Métricas de Avaliação	30
4	RESULTADOS	31
4.1	Classificação com a CNN ResNet-50	31
4.2	Classificação com Associação de Descritores Aprendidos e Wavelet	33
4.3	Considerações Finais	37
5	CONCLUSÃO	39

5.1	Desafios Encontrados	39
5.2	Trabalhos Futuros	40
	REFERÊNCIAS	41

1 Introdução

O câncer é um dos maiores problemas de saúde pública no Brasil e no mundo. Em especial, o câncer de mama é o mais comum entre as mulheres, com 29,7% dos novos casos em 2020, no Brasil. Em consequência do elevado número de casos, a mortalidade por câncer de mama foi de 18.068 óbitos em 2020, cerca de 16,4% do total de mortes por câncer, segundo o Instituto Nacional do Câncer (INCA) (INCA, 2021). Como o câncer de mama se tornou um problema de saúde pública mundial e os melhores índices de um tratamento bem-sucedido ocorrem com diagnóstico da doença em seus estágios iniciais. Por isso, é desejável obter técnicas de diagnóstico mais precisas.

Existem alguns métodos de diagnóstico do câncer de mama, sendo os mais importantes a mamografia, a ressonância magnética e a ultrassonografia. Estes métodos se complementam para a detecção da possível displasia, no entanto, somente a análise do tecido em uma biópsia confirma a existência ou não do tumor. Dentre as técnicas de biópsia mais comuns, a por agulha fina é a mais empregada. O médico patologista então analisa visualmente as células para a realização do diagnóstico por meio de histopatologia (SPANHOL et al., 2016).

A análise dessas imagens histológicas é uma tarefa demorada e depende da experiência do patologista. Para aumentar a acurácia dos diagnósticos, normalmente, é empregado uma dupla leitura da imagem, onde outro especialista faz o diagnóstico no mesmo exame, no entanto, este procedimento aumenta significativamente o custo. Com a possibilidade de armazenamento das imagens com os microscópios eletrônicos, uma alternativa é a utilização de um sistema CAD (do inglês, *Computer Aided Diagnosis*), onde a segunda leitura é feita por um sistema computacional especializado (SPANHOL et al., 2016). O sistema CAD é uma sequência de técnicas de Processamento Digital de Imagens (PDI). Essas técnicas são empregadas nas diversas etapas do sistema desde o melhoramento da imagem até a classificação das lesões. O reconhecimento de padrões e a extração de características das imagens, empregados em um CAD, permitem extrair informações de características morfológicas e não morfológicas (MOREIRA JUNIOR; NASCIMENTO, 2018). Os descritores morfológicos representam as bordas e margens de uma área, enquanto as não morfológicas são as características internas, ou seja, as áreas dentro destas bordas.

Para a identificação de anormalidades em tecidos de mama, diversos algoritmos podem ser empregados. A definição de padrões para os descritores não morfológicos é considerada problemática na interpretação do sistema de visão humana pela dificuldade de interpretação dos níveis de cinza. Por isso, um dos atributos mais importantes para a

representação dessas características são aqueles constituídos pelas informações de textura da imagem.

Uma grande variedade de métodos computacionais pode ser empregado para extração de informações de textura, como matriz de coocorrência, dimensão fractal, filtros baseados em processamento de sinais e muitos outros. Dentre estas alternativas, a transformada *wavelet* vem sendo utilizada em diversos trabalhos na extração de características de imagens histológicas de mama (DUARTE; NASCIMENTO, 2014). Os coeficientes extraídos por esta técnica são utilizados como descritores de características texturais da imagem. Um dos grandes desafios desses descritores é determinar as sub-bandas mais adequadas para permite obter informações de textura que possam representar adequadamente as variações de padrões em tecidos histológicos. Essas técnicas aplicadas para extrair atributos são comumente referidos como “atributos manuais” (do inglês, *handcrafted features*). Uma outra abordagem para a extração de padrões em imagens é a utilização de Redes Neurais Convolucionais (do inglês, *Convolutional Neural Networks* - CNNs). Neste abordagem, o algoritmo através de treinamento, aprende a reconhecer os padrões desejados nas imagens e os descritores produzidos por eles são denominados de "atributos aprendidos"(do inglês, *non-handcrafted features*). Essa abordagem se destaca pela capacidade de arquiteturas de *Deep Learning* de reconhecerem uma gama maior de filtros de padrões complexos em imagens do que os algoritmos especializados (JOYSEEREE et al., 2019).

Uma opção é uma abordagem que permita associar estas técnicas de forma híbrida, em que os coeficientes *wavelet* e os extraídos pela CNN são associados para obter descritores mais precisos. Esta abordagem de associação entre modelos CNN e transformada *wavelet* pode ser encontrada na literatura em estudos que investigaram classificação de texturas (FUJIEDA; TAKAYAMA; HACHISUKA, 2017) e tecido de câncer de pulmão (JOYSEEREE et al., 2019). No entanto, estes não analisaram os níveis de decomposição da *wavelet* na definição de características relevantes para uma classificação. Além disso, não exploraram imagens histológicas em diferentes níveis de magnificação durante a digitalização.

Portanto, a proposta deste trabalho é investigar maneiras de associar as técnicas que permitam descrever informações de textura combinando as CNN e transformada *wavelet*. O qualidade da combinação destes descritores na classificação, será avaliada pela utilização de dois algoritmos de Aprendizado de Máquina (do inglês, *Machine Learning*, ML): o *Support Vector Machine* e o *Random Forest*.

1.1 Objetivos

1.1.1 Geral

Neste projeto pretende-se investigar a associação de descritores *handcrafted* obtidos por meio da transformada *wavelet* e descritores obtidos pela ResNet-50 para classificação de lesões em imagens histológicas obtidas em diferentes magnificações.

1.1.2 Específicos

- Investigar a arquitetura ResNet-50 para extração de características para classificação de lesões da mama;
- Estudar descritores *handcrafted* baseados em imagens de detalhes (vertical, horizontal e diagonal) e aproximação da transformada *wavelet*, *Daubechies-8*, combinados com os descritores da CNN para classificação;
- Investigar a associação entre os descritores *handcrafted* e ResNet-50 para a classificação de lesões de mama;
- Análise de desempenho dos descritores com algoritmos *Support Vector Machines* e *Random Forests* para a classificação.

1.2 Organização desse Documento

Neste capítulo, foram apresentadas as considerações iniciais, motivação e contribuições do desenvolvimento deste projeto. O restante deste trabalho possui a seguinte organização:

- Capítulo 02: Introduz os principais conceitos para a compreensão e validação do trabalho proposto e apresenta os trabalhos correlatos relacionados com a proposta;
- Capítulo 03: Descreve o trabalho proposto, bem como as tecnologias, técnicas e metodologias desenvolvidas;
- Capítulo 04: Apresenta os resultados e as respectivas análises dos modelos de classificação para as lesões de mama.
- Capítulo 05: Apresenta os principais resultados e contribuições obtidas, considerações finais e trabalhos futuros.

2 Contextualização

Neste capítulo são apresentados os conceitos fundamentais para o entendimento deste trabalho.

2.1 Câncer de mama

O câncer pode ser definido como um conjunto de várias doenças que têm em comum o crescimento desordenado de células, que podem invadir tecidos próximos ou distantes, segundo o Instituto Nacional do Câncer ([INCA, 2021](#)). O que diferencia os tipos de câncer são as células de origem, a velocidade de multiplicação e a capacidade de invadir outros tipos de células (metástase).

Uma patologia de câncer pode surgir a partir de causas internas e externas. As causas internas estão relacionadas à genética do indivíduo e a capacidade do organismo em reagir a agressões externas. E as externas estão relacionadas aos hábitos e ao meio sociocultural onde a pessoa está inserida. Esses fatores, sozinhos ou correlacionados, são o que gera a patologia. O fator mais importante no surgimento do câncer é o externo, no entanto, os fatores internos, herdados geneticamente, aumentam em proporções consideráveis a predisposição da doença se desenvolver. De acordo com Seuáñez ([SEUÁNEZ, 2009](#)), “atualmente, estima-se que cerca de 5% a 10% de muitos cânceres estejam associados à predisposição hereditária”.

Em particular, o câncer de mama é o tipo de câncer mais comum entre as mulheres do Brasil e do mundo depois do de pele não melanoma, respondendo por cerca de 28% dos novos casos a cada ano, ([INCA, 2021](#)). Como o câncer de mama é raro em mulheres antes dos 35 anos, a incidência cresce progressivamente, principalmente após os 50 anos. Segundo o INCA, o aumento da incidência de câncer de mama vem crescendo no mundo todo. Embora seja incomum em mulheres jovens, um dos principais fatores de um tratamento de câncer de mama bem sucedido é a detecção precoce. O principal método de detecção precoce, além do exame de palpação, é a mamografia, que pode ser aplicado em mulheres sem sintomas, desde que haja um equilíbrio entre os benefícios e riscos desta prática. No entanto, a maneira mais adequada de verificar a presença de neoplasia no tecido (células cancerígenas) é através da biópsia dos tecidos histológicos ([INCA, 2021](#)). Na Figura 1 são apresentado as áreas da mama, tais como: lóbulos, dutos, etc. Já na Figura 2 é apresentado imagens de tecidos histológicos de mama com lesões do tipo maligna e benigna.

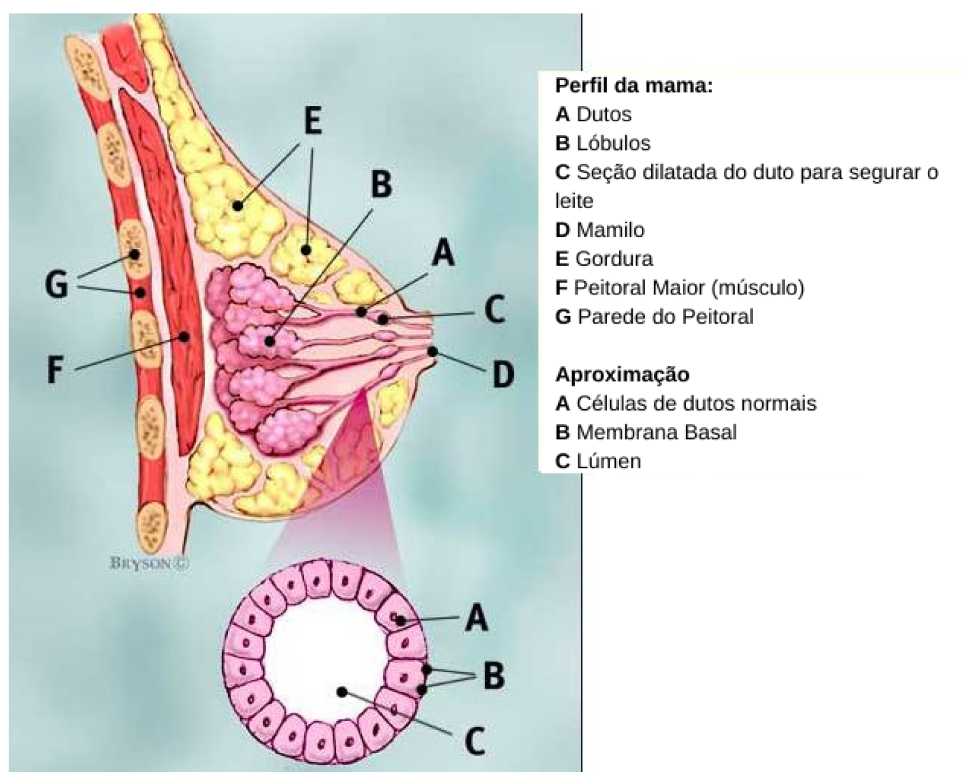
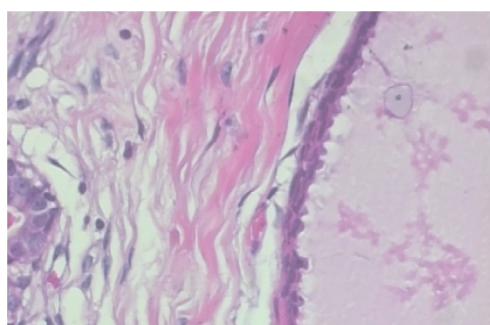
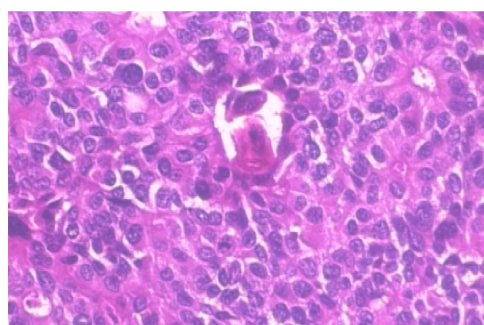


Figura 1 – Anatomia da mama. Fonte: (CANCER, 2021)



(a) Exemplo de lesão benigna.



(b) Exemplo de lesão maligna.

Figura 2 – Exemplos de imagens com lesões de mama benignas e malignas. Fonte: (SPANHOL et al., 2016)

2.2 Sistema de Apoio ao Diagnóstico

Os sistemas CAD, são sistemas computacionais que fornecem informação adicional a um especialista, com o objetivo de melhorar a acurácia do diagnóstico. Uma avaliação computadorizada é útil por ser feita a partir de avaliações objetivas, em contraste com as avaliações por especialistas que estão sujeitos a diversos fatores externos e internos, como a iluminação, cansaço, fadiga visual, distração, entre outros fatores. Em suma, os sistemas CAD podem auxiliar um especialista de duas formas: a primeira, conforme comentado anteriormente, refere-se o auxílio ao diagnóstico, onde o algoritmo extrai características das imagens de entrada e as classifica de acordo com os padrões normais ou anormais para aquele domínio. A outra, consiste no auxílio para detectar lesões e/ou áreas de interesse do especialista (do inglês, *Region Of Interest* - ROI).

Normalmente, os sistemas CADs utilizam técnicas visão computacional e inteligência artificial para produzir seus resultados. Enquanto o primeiro atua nas etapas de processamento das imagens, como ajuste de brilho, realce, diminuição de ruídos, segmentação e extração das características, a inteligência artificial é aplicada à seleção de atributos e no reconhecimento dos padrões para a classificação ou seleção das áreas de interesse. Na Figura 3 pode-se visualizar as etapas que são empregadas num sistema CAD aplicado em imagens histológicas.

2.3 Processamento Digital de Imagens

Uma das subáreas de visão computacional é o processamento de imagens digitais. A evolução da tecnologia dos sistemas computacionais e o desenvolvimento de técnicas no processamento de sinais proporcionaram o desenvolvimento de aplicações como o reconhecimento facial, previsão do tempo, aplicações agrícolas, análise de imagens biomédicas, entre outros.

Os sistemas de processamento de imagens digitais consiste em um *pipeline* composto de cinco etapas: a aquisição da imagem, pré-processamento, segmentação, descrição e reconhecimento.

2.3.1 Aquisição da Imagem

Diversos dispositivos ou sensores podem ser usados para capturar imagens, entre os quais destacam-se as câmeras, os digitalizadores ou *scanners*, aparelhos de radiografia e tomógrafos médicos. Esses dispositivos capturam um sinal analógico e o convertem para uma representação digital.

A representação digital de uma imagem, depende do tipo de sensor utilizado para sua captura, das condições de iluminação e das cores utilizadas, e é armazenada em uma

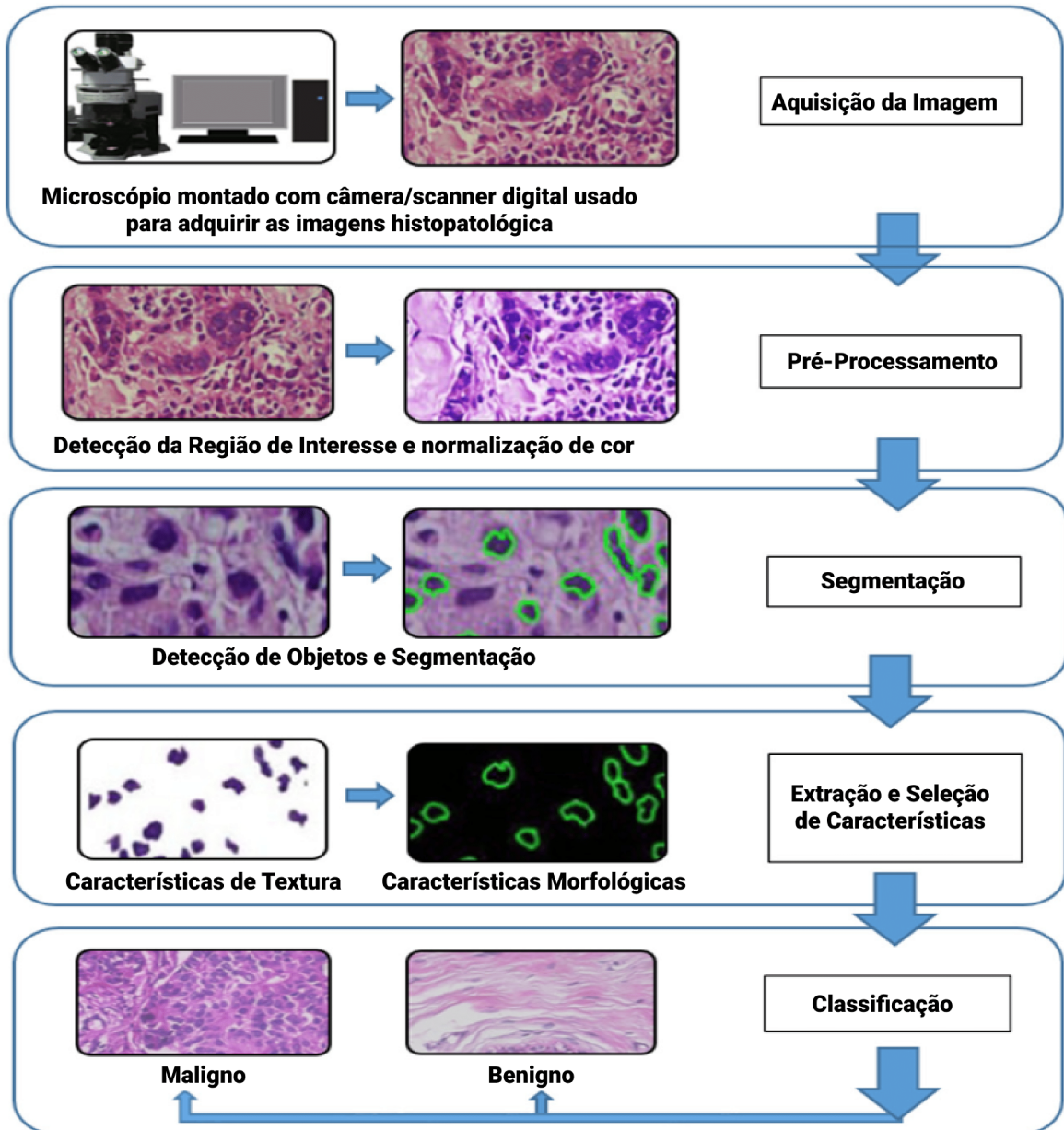


Figura 3 – Etapas empregadas em um sistema CAD - Fonte: (KAUSHAL et al., 2019)

matriz computacional, onde cada elemento determina um ponto da imagem, que também é denominados de pixel. A qualidade e o tamanho da alocação da imagem é diretamente proporcional à quantidade de pixels que ela possui. Uma outra maneira de representar uma imagem, é através de uma função bidimensional $f(x,y)$, em que x e y representam o plano e a amplitude de f em qualquer par de coordenada (x,y) são denominados de nível de cinza ou intensidade da imagem naquele ponto (GONZALEZ; WOODS, 2007).

2.3.2 Pré-Processamento

Durante o processo de conversão do sinal analógico para a representação digital, a imagem digital pode apresentar imperfeições ou degradações, por condições de iluminação ou características internas do dispositivo de captura. Portanto, é necessário aplicar algumas técnicas para atenuar esses efeitos. O agrupamento dessas técnicas é chamado de pré-processamento e tem como objetivo melhorar a qualidade das características da imagem. Essas técnicas se dividem em duas categorias principais: os métodos no domínio espacial, onde os pixels são manipulados diretamente, e os métodos no domínio da frequência, para a modificação da imagem com uma aplicação de uma transformada de Fourier, por exemplo.

2.3.3 Segmentação

Na etapa de segmentação, a imagem passa por um processo de identificação e separação das regiões de interesse. Os principais métodos buscam identificar bordas ou agrupamentos dos valores de intensidade de pixel, separando os objetos do fundo da imagem. Na primeira categoria, os algoritmos identificam mudanças abruptas nos níveis de cinza, que caracterizam as bordas, linhas e/ou descontinuidades. Na segunda abordagem, são identificados os grupos que possuem valores semelhantes para que possam ser agrupados. Geralmente, uma solução adequada é atingida pela combinação de métodos para cada domínio de problema, fazendo com que técnicas específicas tenham preferência sobre as generalistas.

2.3.4 Representação e Descrição

A etapa de representação e descrição refere-se às características ou propriedades que diferenciam os tipos de objetos contidos nas imagens. Na maioria das vezes, essas características são representadas por atributos numéricos, formando um vetor de características que é usado para o reconhecimento de padrões (CONCI; AZEVEDO; LETA, 2021). Esses vetores de características são usados como entradas para os algoritmos de classificação. Por esta razão é crucial coletar informações específicas da imagem em estruturas adequadas de representação, para que os algoritmos de classificação tenham um bom desempenho. As informações extraídas das imagens podem ser agrupadas em dois

grupos: características morfológicas e não morfológicas. As primeiras representam extremidades das imagens como bordas e margens, enquanto as não morfológicas representam informações internas da imagem, como os pixels daquela região.

No domínio das imagens médicas, a extração e representação de características texturais é uma das etapas mais importantes para a identificação das ROIs de um sistema CAD. Por isso, o objetivo da análise de textura é a identificação de atributos ou características que diferenciam o comportamento da superfície. Segundo (GONZALEZ; WOODS, 2007), textura é uma descrição que permite quantificar as suas propriedades de suavidade, rugosidade e regularidade. Na Figura 4 são apresentados exemplos de imagens histológicas de tecido mamário, com diferentes características relacionadas a textura.

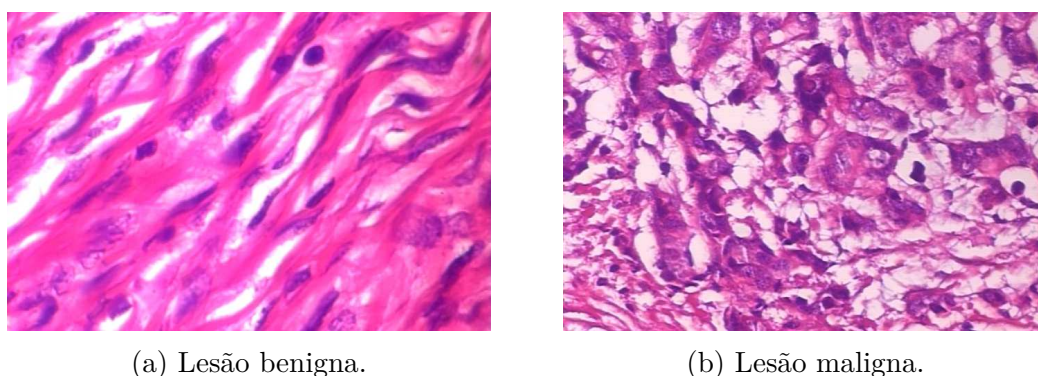


Figura 4 – Exemplos de texturas em uma lesão maligna e uma ROI benigna. Fonte: (SPANHOL et al., 2016)

Em (GONZALEZ; WOODS, 2007), os métodos de extração de características se dividem em três tipos de abordagem:

- Abordagem estatística: uma representação indireta da textura, utilizando propriedades não determinísticas que definem as distribuições e relacionamentos entre os níveis de cinza dos pixels de uma imagem, e que caracteriza a textura como suave, áspera, granulada, entre outros.
- Abordagem estrutural: representa a imagem utilizando texturas primitivas bem definidas (micro texturas) e em arranjos especiais hierárquicos destas primitivas (macro texturas), gerando uma boa descrição simbólica da imagem. Embora seja usada em análise de textura, essa abordagem é mais útil em tarefas onde o objetivo é a síntese de imagens.
- Abordagem espectral: define as representações das imagens, onde os descritores são obtidos a partir da aplicação de uma transformada sobre a imagem. Técnicas como a Transformada de Fourier e Transformada Wavelet são alguns exemplos da abordagem espectral.

Neste trabalho, o método de extração de características será baseado na abordagem espectral, a partir de descritores extraídos de Redes Neurais Artificiais e Transformada Wavelet.

2.3.5 Transformada Wavelet

As transformadas matemáticas são utilizadas nos mais diversos tipos de aplicações, incluindo o processamento e análise de imagens, onde a mais conhecida nesta área é a Transformada de Fourier. Embora seja amplamente utilizada, existem outras técnicas que são mais adequadas para determinados domínios de aplicação, onde outras informações além do tempo e a frequência estão presentes. Em especial, a Transformada Wavelet (TW) é muito utilizada para extração de características espectrais, filtragem, compressão e segmentação de imagens. A ideia por trás da TW é a manipulação de funções para sua representação em outro domínio, que se originou a partir da ideia proposta por Jean Baptise Joseph Fourier.

Em 1822, deu origem à teoria de análise de frequência, a qual demonstra que um sinal periódico pode ser representado pelo somatório de funções sinusoidais, como seno e cosseno. Posteriormente, Jean Fourier demonstrou que esta técnica poderia ser utilizada para descrever qualquer tipo de função. Este método que possibilita a transformação de uma forma complexa em componentes sinusoidais recebeu o nome de Transformada de Fourier (TF).

No entanto, com o passar do tempo, matemáticos começaram a explorar a utilização das TF em diversos domínios de aplicação, como na análise de escala ao invés da análise de frequência. O objetivo era conseguir analisar funções que variam tanto no domínio da frequência quanto no tempo. Assim, surgiu a Transformada por Janela de Fourier (TPJF) ou Transformada de Gabor, criada por Dennis Gabor.

A metodologia da TPJF consiste na criação de uma janela, para delimitar a aplicação da transformada de Fourier apenas na seção definida pelo tamanho da janela. Contudo, a janela possui um valor de largura fixo, o que faz com que ocorra perda de informação temporal do sinal (CONCI; AZEVEDO; LETA, 2021). Para resolver este problema, uma nova técnica foi desenvolvida pelo matemático húngaro Alfred Haar. O método proposto permitia que a largura da janela pudesse variar e assim surgiu a TW. A aplicação de uma Transformada Wavelet permite a criação de janelas com grandes intervalos no tempo para extrair informações de baixa frequência, e intervalos pequenos para a extração das informações de alta frequência (MISITI et al., 2008).

A Transformada Wavelet é representada por um conjunto de funções matemáticas que satisfazem certos requisitos, (OLIVEIRA, 2007) essas funções separam os dados de um sinal em seus diferentes componentes de frequência e extraem cada componente com uma

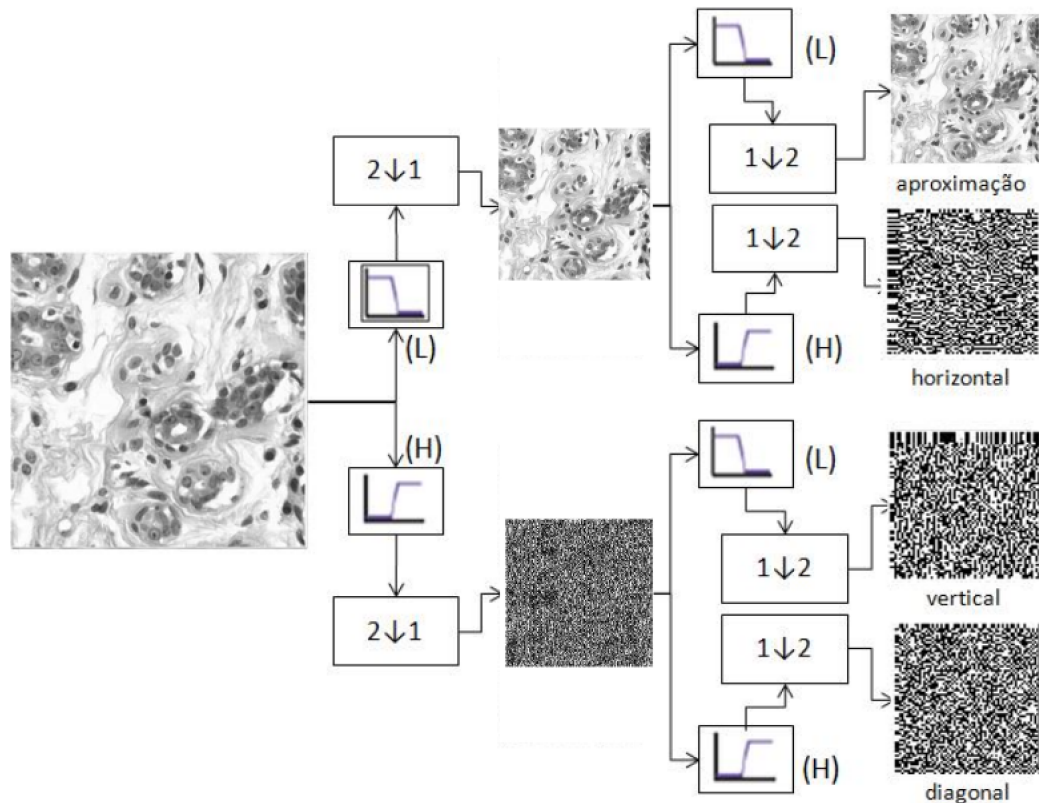


Figura 5 – Aplicação de uma Transformada Wavelet discreta em uma imagem histológica de mama, utilizando a função *wavelet*-mãe Daubechies 4. Fonte: (MOREIRA JUNIOR; NASCIMENTO, 2018)

resolução adequada. Cada função neste conjunto é denominada de função *wavelet*-mãe, sendo a Haar, Daubechies e Symlet as mais utilizadas para o processamento de imagens digitais. Existem dois tipos de Transformas Wavelet as discreta e a contínua. As contínuas produzem muitas informações redundantes, com muitas informações desnecessárias para o processamento de uma imagem, mas que são úteis em processos como eliminação de ruídos em sinais e compressão de dados.

A aplicação da Transformada Wavelet é usada como uma espécie de filtro nos problemas de visão computacional, que quando aplicada com uma janela grande (filtro H) extrai informações mais genéricas sobre a imagem e quando aplicada em uma janela pequena (filtro L), extrai informações locais de uma determinada região da imagem. A combinação desses filtros produz quatro coeficientes, denominados também de sub-bandas. Quando aplicado os filtros LL, são produzido os coeficientes de aproximação, os filtros LH produzem os coeficientes de detalhe horizontal, os filtros HL os coeficientes de detalhe vertical e os filtros HH os coeficientes de detalhe diagonal, conforme exemplificados na figura 5.

2.4 Redes Neurais Convolucionais

Após a extração dos descritores na última etapa do processamento de imagens, a sequência natural é a classificação de cada conjunto de informações em uma classe ou intervalo de valores que o represente. As principais técnicas de classificação, são baseadas em algoritmos de ML. Nesses algoritmos, o modelo recebe um conjunto de dados, neste caso imagens, que ele utiliza para calibrar seus parâmetros internos para “aprender” a identificar a qual classe cada instância do conjunto de dados pertence.

Os algoritmos de aprendizado de máquina são uma subárea da Inteligência Artificial (do inglês, *Artificial Intelligence* - AI), um dos mais novos campos de conhecimento em ciência e engenharia. Apesar do desejo de criar máquinas capazes de pensar ser tão antigo, os avanços na área acompanharam a criação e o desenvolvimento dos computadores. Os primeiros trabalhos em IA, datam do final da segunda guerra mundial, com o trabalho de Warren McCulloch e Walter Pitts em 1943. Eles propuseram um modelo de neurônios artificiais, onde cada neurônio poderia estar em dois estados possíveis: ligado ou desligado. A mudança para o estado ligado dependia do estímulo de uma quantidade mínima de vizinhos também ligados. Eles mostraram que a combinação de neurônios poderia ser utilizada para a criação de estruturas lógicas, como *E*, *OU* e *NÃO*. Eles também sugeriram que essas combinações de neurônios, denominadas Redes Neurais Artificiais (RNAs) poderiam aprender.

As RNAs, fazem parte das técnicas de aprendizado de máquina e podem ser usadas em tarefas de classificação, aproximação de funções, entre outros. As RNA são “um sistema composto por um número de elementos de processamento simples, altamente interligados, que processam a informação pela sua resposta de estado dinâmica a entradas externas.” (CAUDILL, 1989). Essa técnica é composta por componentes simples denominados neurônios artificiais, que são funções não-lineares simples, que quando interligadas, simulam a interação dos neurônios no sistema nervoso central. Na Figura 6 é ilustrada a diferença entre neurônios artificiais e reais.

Normalmente, os neurônios artificiais que possuem tarefas semelhantes são agrupados em camadas. Essas camadas são denominadas ‘camada de entrada’, ‘camada oculta’ e ‘camada de saída’. A primeira é responsável por receber os dados a serem interpretados pela rede e a última classifica a informação que foi recebida na entrada. A existência da camada oculta é opcional, no entanto, as RNAs sem camadas ocultas resolvem apenas problemas que são linearmente separáveis. O conjunto de camadas ocultas produzem abstrações que interpretam os dados recebidos pela camada de entrada e permitem a resolução de problemas mais complexos.

O ponto chave dos algoritmos de aprendizado de máquina, incluindo as RNAs, é a capacidade de “aprender” através de treinamento. Para as tarefas de classificação, a

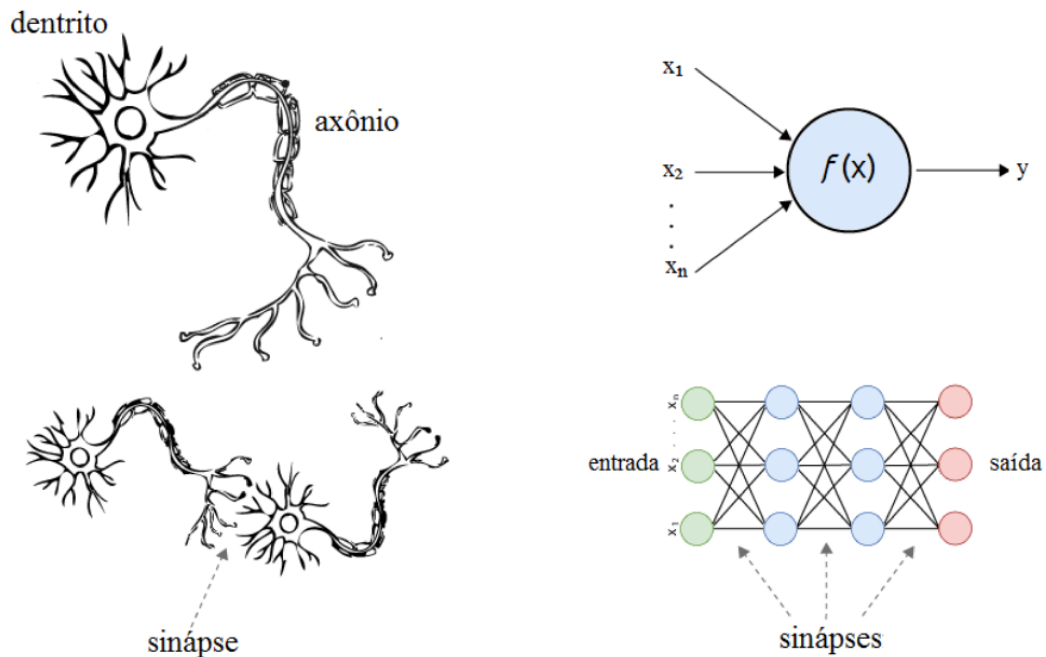


Figura 6 – Representação de neurônios reais (esquerda) e artificiais (direita). Fonte: (ZUIN; VELOSO, 2017)

técnica utilizada é o aprendizado supervisionado, no qual um conjunto de dados é separado em amostras de treinamento e validação. As amostras de treinamento são divididas em pequenos conjuntos denominados *batches*, que são inseridos no modelo, que tenta prever de qual classe cada um dos itens pertencentes àquele *batch*. Neste momento, uma função de custo é aplicada para calcular a quantidade de amostras que o modelo classificou corretamente e a partir deste resultado é função de aprendizado atualiza os pesos entre as conexões da rede. Após todos os *batches* do conjunto de treinamento serem utilizados, testa-se o conjunto de validação (amostras que são desconhecidas a RNA) para simular o seu comportamento em cenários desconhecidos, ou seja, se o modelo aprendeu a classificar de forma generalizada, ou somente aprendeu a identificar as características do conjunto de treinamento, em um problema chamado de *overfitting*.

Com o passar do tempo, diversas arquiteturas de RNAs foram surgindo e se especializando em problemas específicos. Para o processamento de imagem, uma das abordagens mais bem sucedidas são as CNN. Elas foram introduzidas por (AL, 1989) e são modelos que usam uma operação chamada de convolução, que visa reduzir e transformar as dimensões de imagem, gerando uma versão simplificada no entanto representativa (GO-ODFELLOW; BENGIO; COURVILLE, 2016).

A operação de convolução é realizada através dos filtros convolucionais, que são matrizes de dimensão $X \times Y \times Z$, e percorre toda a imagem de entrada, representada por uma matriz de pixels. A cada etapa do deslocamento, cada valor do filtro de deslocamento é multiplicado pelo seu respectivo na matriz de representação. O resultado dessa opera-

das camadas mais rasas nas mais profundas, permitindo redes com mais de 100 camadas de profundidade. Os detalhes de funcionamento da ResNet serão apresentados na seção da metodologia.

2.5 Trabalhos Correlatos

Nesta seção, apresentamos algumas das principais contribuições na literatura que estão relacionadas com o trabalho proposto para imagens histológicas de mama.

O BreakHis dataset (SPANHOL et al., 2016) é um conjunto de dados que provê imagens histológicas de lesões de mama, com quatro níveis de magnificação (40x, 100x, 200x, 400x), contendo amostras de tumores benignos e malignos. Em relação a esta base de imagens, no trabalho de (DIMITROPOULOS et al., 2017), foi apresentada uma técnica para identificar lesões a partir de um VLAD (*Vector of Locally Aggregated Descriptors*), que é um conjunto de sinais multidimensionais que são representados como uma nuvem de pontos Grassmannianos, que representam a dinâmica e a aparência da imagem. Utilizando das propriedades geométricas da nuvem de pontos, estima-se a codificação de VLAD (DIMITROPOULOS et al., 2017), para classificar o carcinoma invasivo. Esta abordagem, avaliou os diferentes níveis de magnificação separadamente, em patches de 8x8 pixels, 16x16 pixels, 32x32 pixels e 64x64 pixels, e o conjunto de dados foi subdividido em 70% para treinamento e 30% para validação do modelo, utilizando o algoritmo de aprendizado de máquina SVM. O melhor resultado foi atingido utilizando os *patches* 8x8 pixels com 100x de aproximação, resultando em 92,1% de acurácia.

Os mesmos autores da base de imagens BreakHis, analisaram algumas arquiteturas de aprendizado profundo para as imagens da base de dados (SPANHOL et al., 2016). Neste estudo, eles utilizaram o modelo da LeNet e da AlexNet, com *patches* de 32x32 e 64x64 com todos os níveis de magnificação. A LeNet alcançou uma acurácia de 72%, enquanto a AlexNet conseguiu 84% de acurácia em média, com 200x de magnificação.

Em (BAYRAMOGLU; KANNALA; HEIKKILÄ, 2016), os autores propuseram um método de classificação baseado em CNN. O modelo proposto, realizava classificação binária no modo de classificação único ou além da classificação, o modelo identifica o nível de magnificação da imagem inserida num modelo de classificação multitarefa. A rede possuía três camadas convolucionais, em que cada uma era seguida de uma operação de ReLU (*Rectified Linear Unit*) e uma camada de *pooling*. Foi realizado um processo de aumento dos dados, através de rotação e um *crop* no centro da imagem para criar um *patch* quadrado. Como muitas vezes tarefas diferentes compartilham um conjunto de características semelhantes, para a criação de uma rede multitarefa, a arquitetura da rede foi modificada para gerar duas saídas na última camada antes da classificação os quais eram inseridos em duas camadas *softmax* paralelas. O modelo proposto atingiu 83,25%

de acurácia média, no modo de tarefa única e uma taxa de acurácia média de 82,13% para classificação entre benigno e maligno e 80,10% na classificação dos níveis de magnificação. Os autores afirmam que redes mais profundas poderiam ser investigadas, assim como o aumento de imagens histológicas de mama podem melhorar a acurácia de classificação nos modelos baseados em CNN.

No trabalho de (HAN et al., 2017), os autores propuseram um modelo de aprendizado profundo baseado em CNN, para a multi-classificação de imagens histopatológicas de mama. Essa abordagem explorou classificar o tipo de câncer, ao invés da classificação binária - maligno e benigno. Esta nova técnica proposta, foi uma CNN baseada em estruturas de classe (do inglês, *Class Structure-Based Deep Convolutional Neural Networks - CSDCNN*), que é um modelo de aprendizado de representação não-linear que adota um treinamento fim a fim de uma maneira que ele aprende automaticamente a semântica e as características hierárquicas discriminativas. O modelo atingiu uma acurácia média de 93,2% utilizando transferência de aprendizado para afinação dos parâmetros da rede.

2.6 Considerações Finais

Conforme exemplificamos anteriormente, o câncer de mama é um problema relevante no mundo todo, por isso são desejáveis modelos que auxiliem no diagnóstico preciso nos estágios iniciais da doença, onde o tratamento é mais efetivo e menos agressivo para o paciente. Em relação ao banco de imagens BreakHis, devido a complexidade das imagens e da não-homogeneidade das características de tipos diferentes de câncer, tornam os modelos de aprendizado profundo interessantes para a extração de características aprendidas (*non-handcrafted*).

Para esta base de dados, os principais trabalhos abordaram o uso de CNN para seus resultados. No entanto, notou-se que o uso da associação de características *handcrafted* e *non-handcrafted* não foram explorados. Por isso, este trabalho busca explorar a combinação dos vetores de características extraídos por uma rede neural profunda e as informações produzidas utilizando a TW, que são uma das técnicas mais importantes para a extração de características *handcrafted* de imagens histopatológicas, para alcançar uma melhor acurácia média na identificação de lesões de mama em sistemas CAD.

3 Metodologia

Neste capítulo são apresentadas as informações sobre o banco de imagens, as técnicas desenvolvidas para associação das informações das características, assim como os classificadores empregados para avaliação dessas características para discretizar as classes de lesões.

3.1 Base de Dados

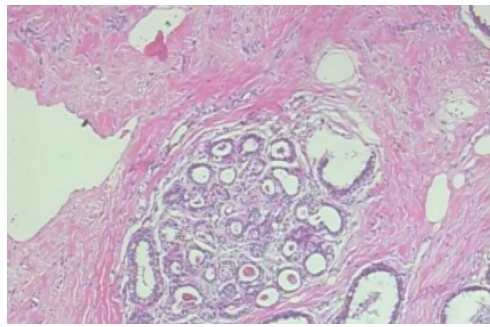
O BreakHis dataset, é uma base de dados composta de imagens microscópicas de tumores de mama benignos e malignos, coletadas de 82 pacientes pelo Laboratório de P & D - Anatomia Patológica e Citopatologia, da Universidade Federal do Paraná, Brasil, entre Janeiro e Dezembro de 2014. As imagens foram geradas a partir de amostras de tecido mamário em lâminas de biópsia, extraídos por mastectomia parcial ou biópsia excisional, que é um procedimento cirúrgico que remove uma amostra maior de tecido, do que os métodos convencionais de biópsia por agulha e é realizado em um hospital com anestesia geral (SPANHOL et al., 2016).

O conjunto de dados é composto de 9.109 imagens, contendo 2.480 amostras benignas e 5.429 malignas, com 4 tipos de lesões diferentes em cada subconjunto: adenosis, fibroadenoma, filodes e adenoma tubular no grupo benigno, e carcinoma ductal, carcinoma lobular, carcinoma mucinoso e carcinoma papilar no grupo maligno. As imagens foram extraídas usando diferentes níveis de magnificação (40x, 100x, 200x e 400x) com uma resolução de 700 x 460 pixels em formato *PNG* com 3 canais de cores RGB.

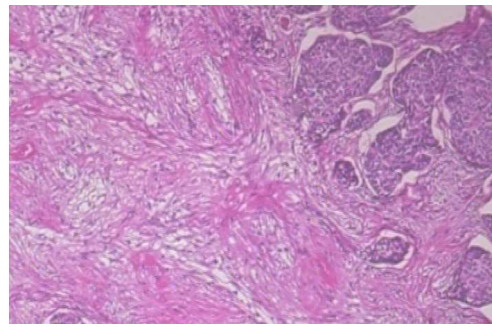
As dimensões das imagens originais são de 700x460 pixels, realizou-se um redimensionamento para utilizá-las como entrada do modelo de Aprendizado de Máquina. Redimensionou-se então, as imagens para 224x224 pixels, para corresponder com a entrada esperada do modelo da CNN. Foram utilizadas 1.370 imagens malignas e 598 benignas com magnificação de 40x, 1.437 malignas e 614 benignas com magnificação de 100x, 1.390 malignas e 594 benignas com magnificação de 200x e 1.232 malignas e 562 benignas com magnificação de 400x. Na Figura 8, são exibidas algumas imagens com diferentes níveis de magnificação do conjunto de dados.

3.2 Modelo Proposto

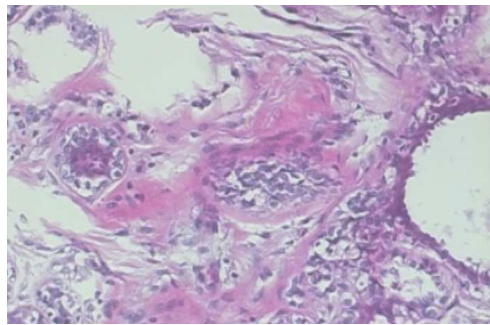
O modelo proposto foi composto pela arquitetura *Deep Residual Network* (ResNet), com 50 camadas de profundidade (ResNet-50) para a extração de características



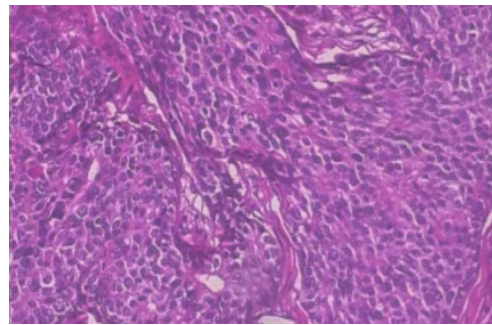
(a) Lesão benigna com magnificação de 40x.



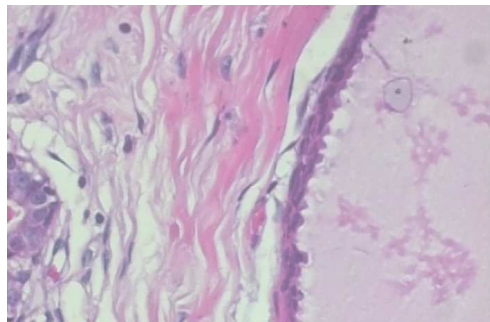
(b) Lesão maligna com magnificação de 40x.



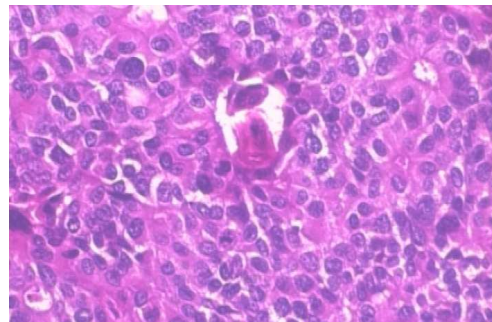
(c) Lesão benigna com magnificação de 100x.



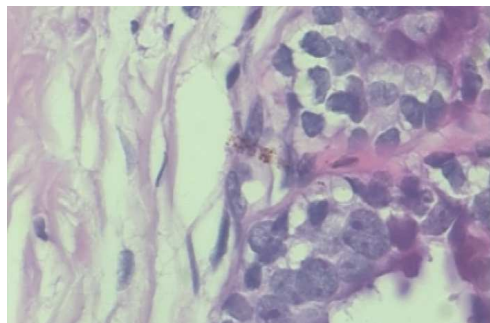
(d) Lesão maligna com magnificação de 100x.



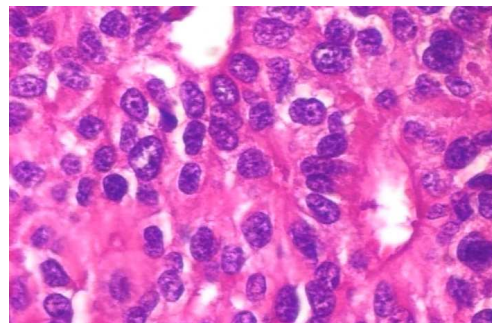
(e) Lesão benigna com magnificação de 200x.



(f) Lesão maligna com magnificação de 200x.



(g) Lesão benigna com magnificação de 400x.



(h) Lesão maligna com magnificação de 400x.

Figura 8 – Exemplos de imagens benignas e malignas com múltiplos fatores de magnificação. Fonte: (SPANHOL et al., 2016)

no – *handcrafted* na camada de *flatten*, os quais foram associados com os descritores *handcrafted* extraídos pela aplicação da Transformada Wavelet. Os descritores extraí-

dos foram associados das seguintes maneiras: i) concatenação do vetor de características da *ResNet-50* e as sub bandas de detalhe (vertical, horizontal e diagonal) extraídos da aplicação da *Wavelet*; ii) concatenação do vetor de características da *ResNet-50* com o descritor energia de cada subbanda de detalhe (vertical, horizontal e diagonal) extraídas da *Wavelet*; iii) concatenação do vetor de características da *ResNet-50* com a sub-banda de aproximação extraída da *Wavelet*; iv) concatenação do vetor de características da *ResNet-50* com a energia da subbanda de aproximação extraída da *Wavelet*.

Os vetores resultantes dessas associações de características foram classificados por modelos de aprendizado de máquina, o *Support Vector Machine* (SVM) e o *Random Forest* (RF). Na Figura 9 é exemplificada a abordagem proposta.

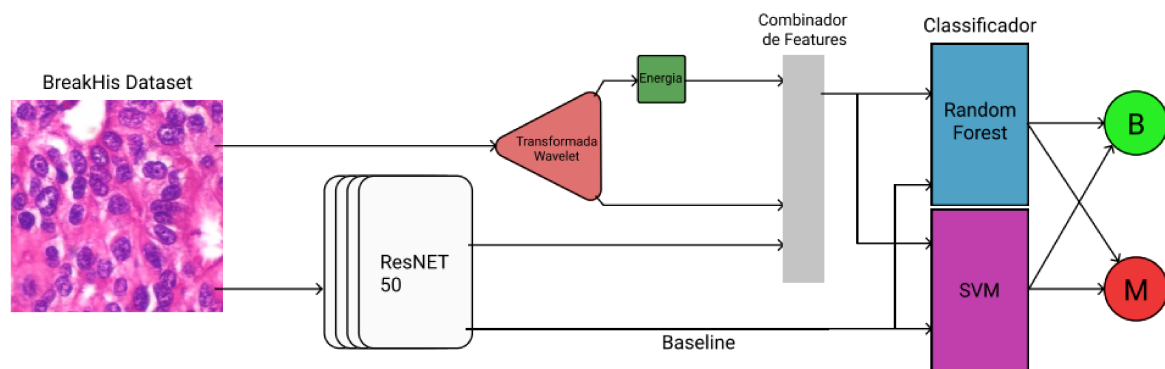


Figura 9 – O modelo proposto combina os descritores *handcrafted* extraídos pela *wavelet*, com os descritores *non-handcrafted* extraídos pela *ResNet-50*.

3.3 Descritores

Os descritores são objetos compostos por informações sintetizadas descritivas da imagem, e a utilização de técnicas que produzam descritores de qualidade são de suma importância para determinar o grau de sucesso da solução. Essas informações, também denominadas mapa de características, são a representação condensada do conteúdo da imagem, as quais são interpretadas para associá-lo a classe à qual pertence. Os mapas de características extraídos das imagens pode ser subdivido de acordo com o tipo de método utilizado para a sua extração. Os que utilizam métodos feitos a mão, especializados à um problema específico, produz descritores denominados *handcrafted*. Os descritores produzidos por métodos generalistas, são denominados *non-handcrafted*. Neste trabalho, apresentou-se uma proposta que utilize a associação dos descritores *handcrafted* e *non-handcrafted* para a classificação de lesões histopatológicas de mama. Foram utilizados descritores *non-handcrafted* gerados por uma arquitetura *ResNet-50*, associados com des-

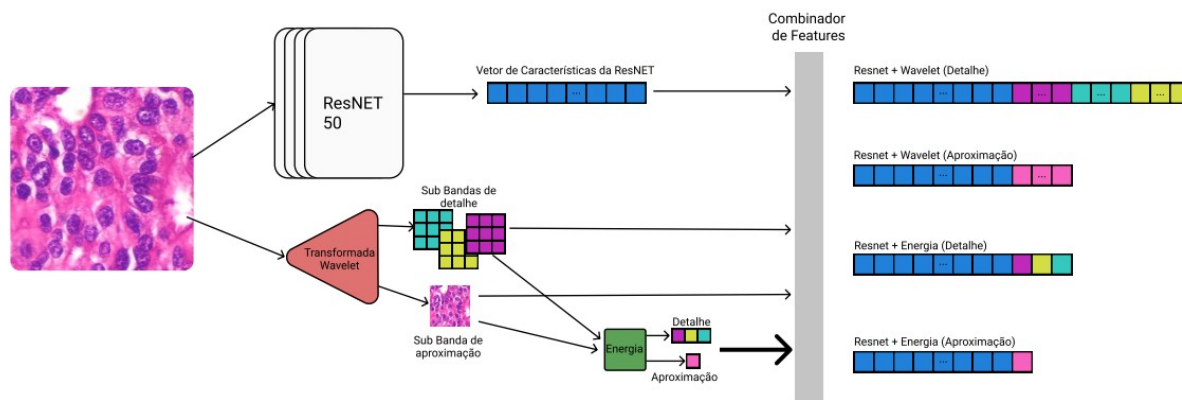


Figura 10 – Representação das combinações de *features* produzidas no processo de classificação.

critores *handcrafted* produzidos pela aplicação da transformada *Wavelet* com função mãe *Daubechies-8*.

Os descritores *non-handcrafted* são extraídos da camada de *flatten* da *ResNet-50* e são representados por um vetor numérico. Para calcular os coeficientes descritores *handcrafted*, cada imagem correspondente ao vetor de características extraído pela *ResNet-50*, foi convertida para uma imagem em nível de cinza, com um único canal de cor, e aplicou-se se transformada *Wavelet Daubechies-8*. A aplicação da TW produz quatro matrizes de coeficientes, que são agrupadas em duas subclasses: as sub-bandas de detalhe (vertical, horizontal e diagonal) e a sub-banda de aproximação. Então, cada matriz de sub-banda foi transformada em um vetor e associado ao vetor extraído da *ResNet-50*. Na Figura 10 é apresentado o processo de associação dessas informações para a classificação das lesões.

3.4 ResNet-50

Com a expansão da utilização das CNNs, acreditava-se que redes mais profundas seriam a solução para atacar problemas de maior complexidade. Os autores em (HE et al., 2016), introduziram um novo conceito na criação de camadas em CNNs, que são as *shortcut connections*. Nesta nova técnica, a saída de uma camada é usada como entrada da próxima camada convolucional e também é usada como entrada em uma camada mais profunda, como se a informação pulasse algumas camadas para ser processada por camadas mais profundas. Desta forma, o problema do desaparecimento do gradiente descendente foi mitigado, permitindo o treinamento de arquiteturas muito profundas (1000 camadas). A partir desta nova operação, criou-se as arquiteturas de Redes Convolucionais Residuais, *ResNet*, que são extensamente utilizadas para problemas de visão computacional complexos. A Figura 11 exemplifica como as *shortcut connections* são utilizadas em um modelo

ResNet.

Apesar do problema do gradiente descendente ter sido minimizado, treinar redes profundas é uma tarefa muito sensível. A quantidade de dados necessários para o treinamento acompanhou proporcionalmente à profundidade das camadas. Por isso para gerar modelos robustos, dezenas de milhares de imagens são necessárias. Por esta razão, outras alternativas de treinamento foram propostas, uma delas é a transferência de aprendizado, que consiste em usar uma rede robusta, treinada previamente com um grande conjunto de dados. A transferência de aprendizado é uma estratégia de treinamento que consiste no uso de pesos de uma CNN treinada previamente. Então, foi empregada a estratégia de transferência de aprendizado, inicializando o modelo com os pesos usados na competição *ImageNET* (DENG et al., 2009).

Neste trabalho, utilizou-se o modelo *ResNet-50*, que é uma implementação da arquitetura *ResNet* com 50 camadas de profundidade, disponível na biblioteca *TensorFlow* (ABADI et al., 2015). Para o treinamento, empregou-se um **Batch Size** de 32, e o conjunto de dados foi dividido em 70% das imagens para treinamento e 30% para validação. Utilizou-se o otimizador Adam (KINGMA; BA, 2017), com uma taxa de aprendizado, *Learning-Rate*, de 0.5×10^{-6} , com 10 épocas de treinamento.

3.5 Transformada *Wavelet*

Para a extração das características baseadas na transformada *Wavelet*, utilizou-se a biblioteca PyWavelet (LEE et al., 2019), com um único nível de decomposição, aplicados às imagens em escala de cinza. Escolheu-se a Transformada *Wavelet* discreta com função-mãe *Daubechies 8*, pois de acordo com (NASCIMENTO et al., 2013), a utilização da *Wavelet*-mãe *Daubechies-8* produz descritores mais representativos para as imagens histológicas de mama. A aplicação da Transformada *Wavelet*, produziu quatro sub-bandas (imagens): aproximação, detalhe vertical, detalhe horizontal e detalhe diagonal.

Também foi avaliado diminuir a dimensionalidade dos descritores *handcrafted*, o calculo apenas da energia sobre as sub-bandas. Para isso, empregou-se:

$$Energia(D_j) = \sqrt{\frac{1}{N_1} \sum_{i=1}^{N_1} (D_j)^2[i]}$$

,

onde D_j é o vetor de detalhe, j indica qual sub-banda de detalhe e N_1 é o tamanho do vetor naquele nível de decomposição.

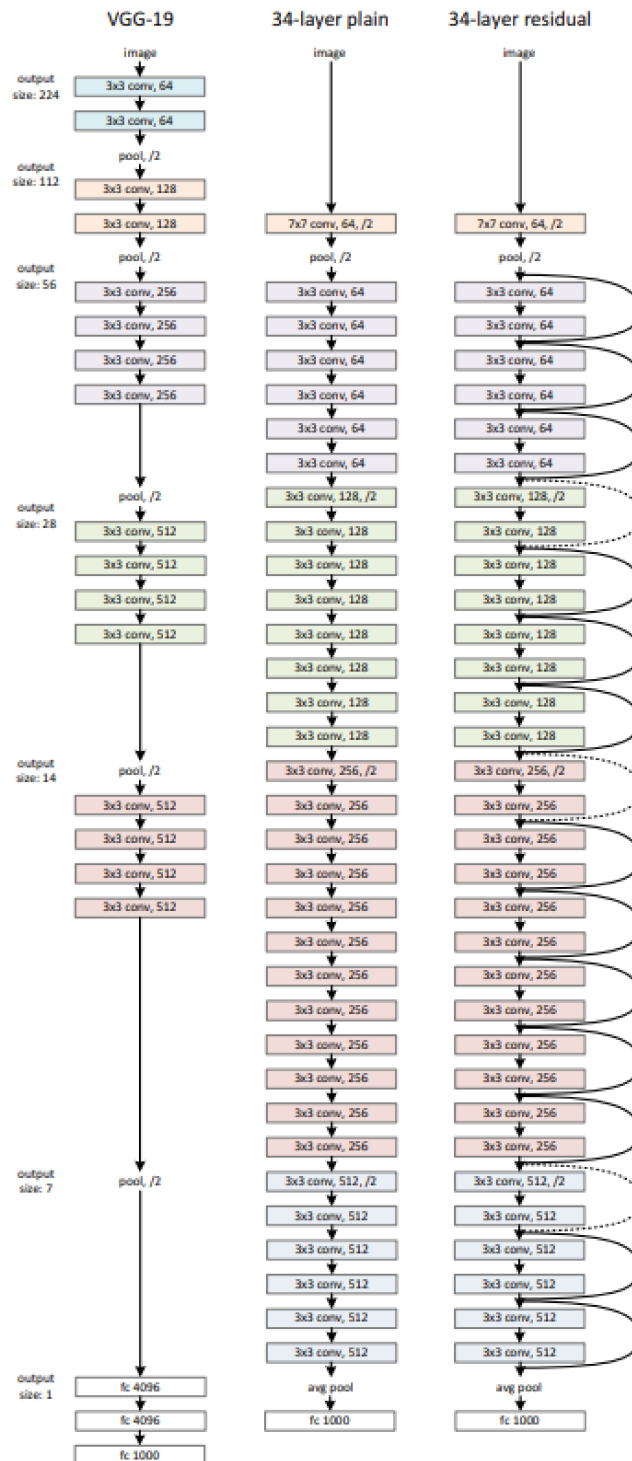


Figura 11 – A imagem mostra alguns exemplos de arquiteturas profundas, como a VGG (SIMONYAN; ZISSERMAN, 2014), e outras redes com e sem utilização de resíduo.

3.6 Classificação

Para classificar a combinação dos descritores e compará-los com o desempenho da *ResNet-50*, foram empregados dois dos algoritmos de aprendizado de máquina mais utilizados na literatura. As técnicas escolhidas, foram o Random Forest e o Support Vector Machines, pela sua capacidade de obter resultados com desempenho do estado-da-arte nos mais diversos domínios de aplicação.

3.6.1 Random Forest

O classificador Random Forest (RF) é um algoritmo baseado em árvores de decisão, onde cada árvore cresce de acordo com coeficientes aleatórios. A principal característica deste algoritmo, é sua capacidade de processar grandes quantidades de dados de maneira robusta e eficiente na etapa de treinamento baseado nas árvores de decisão (KO; KIM; NAM, 2011). Segundo (RAMOS; NASCIMENTO; PEREIRA, 2012), são utilizados quatro etapas para a construção de cada árvore de decisão:

- Etapa 1 - obtenção dos dados por um processo de amostragem aleatória, onde $\frac{2}{3}$ dos dados são utilizados para construir a árvore;
- Etapa 2 - são selecionados um número aleatório de atributos do conjunto de descritores. Desse conjunto, é selecionado o atributo que possuir o maior número de informações e é realizada uma comparação em cada nó;
- Etapa 3 - o processo de seleção da etapa anterior é realizado até que nenhum novo nó possa ser criado;
- Etapa 4 - a partir do $\frac{1}{3}$ dos dados separados na etapa 1, verifica-se o erro e a correção é calculada.

3.6.2 Support Vector Machine

O Support Vector Machine (SVM) é uma técnica de classificação baseada em aprendizagem estatística. Assumindo que os dados são linearmente separáveis, esta técnica encontra um hiperplano com separação máxima entre as classes (EVGENIOU; PONTIL, 2001). O algoritmo utiliza uma função de *kernel* para mapear os dados para uma dimensão maior, até que os dados se tornem linearmente separáveis. Quando o algoritmo para os dados estão separados em dois hiperplanos paralelos e quanto maior for a distância entre estes planos, menor será a taxa de erro na classificação.

Neste trabalho, a implementação de RF e do SVM foi realizada por meio da biblioteca Scikit-Learn para Python (PEDREGOSA et al., 2011). Foram utilizados as configurações já estabelecidas na biblioteca.

3.7 Métricas de Avaliação

Para avaliar o desempenho do modelo, foram utilizados as seguintes métricas: acurácia e área sob da curva ROC (do inglês, *Area Under the ROC Curve* - AUC). A acurácia é taxa de acertos dividida pela quantidade total de amostras testadas. Essa é um bom indicador de desempenho geral, no entanto pode ser afetada pelo desbalanceamento do conjunto de dados.

A métrica AUC é utilizada para calcular o desempenho do modelo em classificar corretamente classes positivas como positivas e classes negativas como negativas. Esta métrica é menos sensível ao desbalanceamento no conjunto de dados, o que é uma bom indicativo em problemas onde a sensibilidade da classificação é importante. Ambas as métricas foram calculadas utilizando a implementação da biblioteca Scikit-Learn (PE-DREGOSA et al., 2011).

4 Resultados

Neste capítulo, são apresentados os resultados obtidos com a metodologia proposta. Primeiro, apresentado os resultados denominado *baseline*, que consiste na classificação das *features* extraídas pela ResNet-50 sem uso de associação de descritores. Posteriormente, são tratados os resultados da associação dos descritores *handcrafted* e *non-handcrafted*.

4.1 Classificação com a CNN ResNet-50

Na primeira etapa foi realizado o treinamento da ResNet-50, a qual foi empregada na classificação e na extração de características das imagens. O treinamento foi realizado utilizando 10 épocas. Nas Figuras 12(a), 12(b), 12(c) e 12(d) são apresentadas as informações a respeito do desempenho da CNN ResNet-50 na fase de treinamento e validação para cada conjunto imagens histológicas com diferentes magnificações. Esses gráficos permitem mostrar o desempenho do processo de aprendizado da rede e também as características relacionadas à taxa de erro. É possível perceber que em todas as magnificações, a acurácia do treinamento foi de quase 100%, no entanto, a mesma acurácia não foi alcançada para a validação. É possível notar que houve um bom desempenho dos valores com a métrica AUC em todas as magnificações, tanto em treinamento quanto em validação.

Para avaliar as características extraídas dos modelos, foi realizado o teste estatístico de Mann-Whitney com uma função Função de Distribuição Empírica Acumulativa (do inglês, *Empirical Cumulative Distribution Function* - ECDF). O teste de Mann-Whitney tem o objetivo de definir se os diferentes grupos de dados estudados pertencem ou não à mesma população. A partir do teste é calculado um p -valor, que representa a probabilidade da hipótese ser nula. Nesse caso, a hipótese de que características extraídas são separáveis é investigada, isto é, pertencem a grupos diferentes. Quanto menor for o p -valor melhor o resultado, respeitando sempre o intervalo de confiança de 95%.

Na Figura 13 é possível observar os resultados após aplicação deste teste sobre os descritores aprendidos na camada de *flatten* da CNN. Nesses dados, pode ser observado que acima de 50% das amostras possuem um p -valor menor ou igual a 0,05, ou seja, são separáveis com um intervalo alto de confiança. A outra metade das amostras, possui um crescimento de p -valor linear, indicando que o conjunto de dados não possui uma boa forma de separação.

Nas Tabelas 1 e 2 são apresentados os resultados da classificação da rede ResNet-50 e usando apenas as características aprendidas na camada de *flattting* e associadas aos algoritmos de aprendizado de máquina. Essas tabelas são denominadas o *baseline* para as investigações propostas nesse estudo. Nesta avaliação observa-se que os valores da

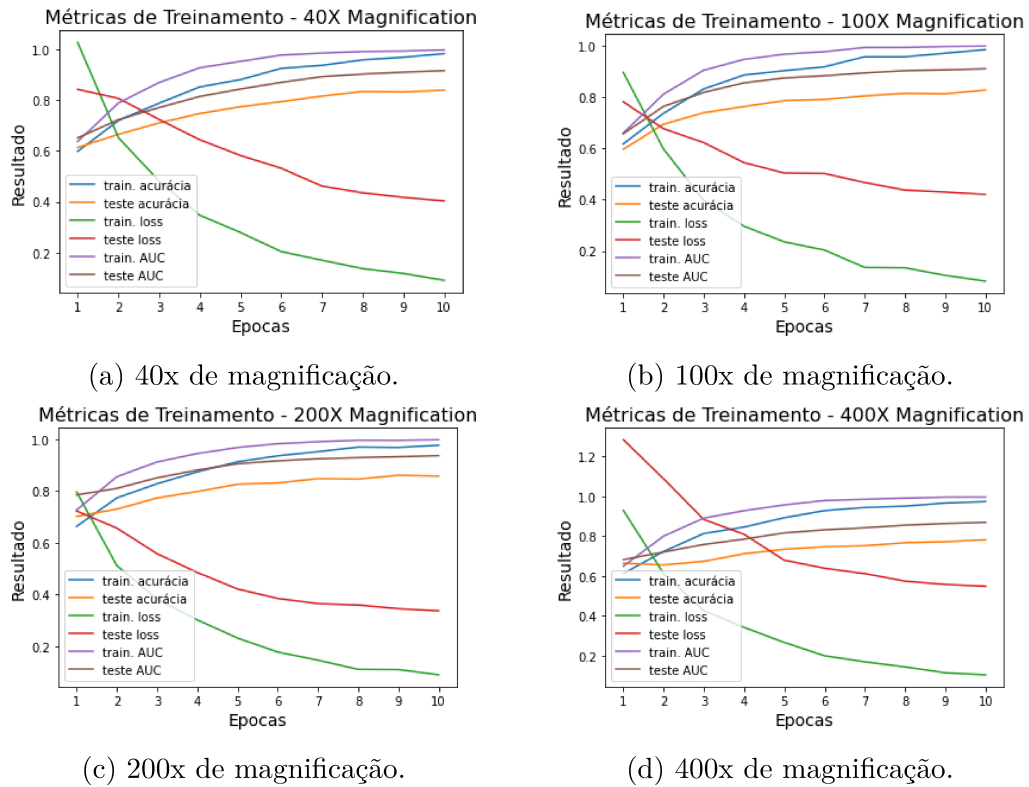


Figura 12 – Avaliação do processo de treinamento e teste na arquitetura ResNet-50 com diferentes níveis de magnificação das imagens histológicas.

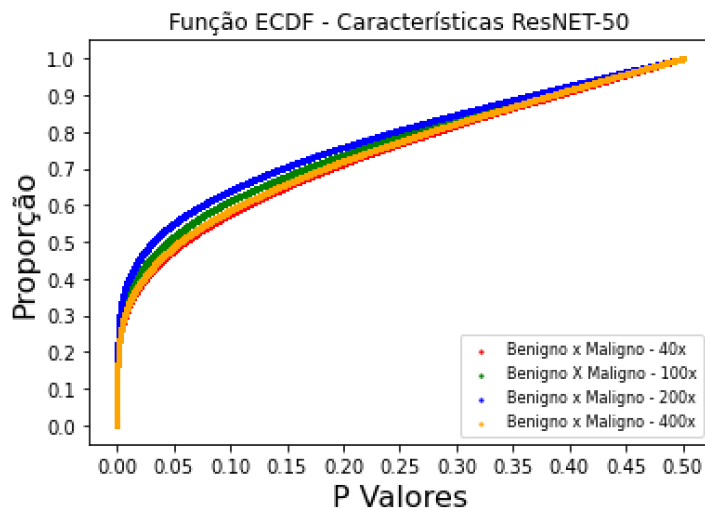


Figura 13 – ECDF para as características extraídas do modelo ResNet-50.

acurácia foram inferiores a 90% em todos os níveis de magnificação e que a SVM produziu as melhores taxas de acurácia, com 89,5% para as imagens com 200x de magnificação. Em contra partida, o classificador Softmax presente na última camada da ResNet-50, atingiu os melhores resultados em AUC.

Tabela 1 – Resultados obtidos com a métrica acurácia para as características extraídas com a ResNet-50.

Classificador	40X	100X	200X	400X
Camada <i>Softmax</i>	83,9%	82,7%	85,7%	78,2%
RF	81,4%	81,0%	88,0%	79,8%
SVM	88,4%	88,4%	89,5%	83,1%

Tabela 2 – Avaliação da métrica AUC para as características extraídas da ResNet-50.

Classificador	40X	100X	200X	400X
Camada <i>Softmax</i>	0,916	0,910	0,936	0,870
RF	0,717	0,712	0,807	0,711
SVM	0,752	0,830	0,840	0,760

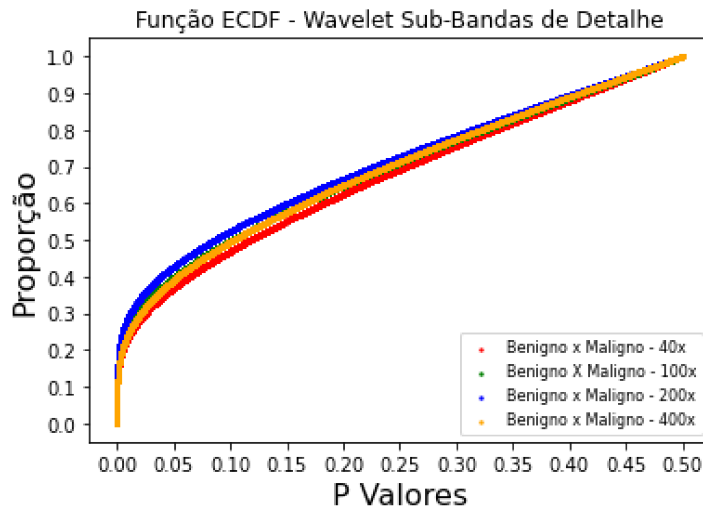


Figura 14 – ECDF para as características extraída do modelo ResNet-50 combinadas com as sub-bandas de detalhe da transformada wavelet.

4.2 Classificação com Associação de Descritores Aprendidos e Wavelet

Após a avaliação das características *non-handcrafted*, a etapa seguinte consiste na combinação com os descritores *handcrafted* a partir da aplicação da transformada Wavelet com a função-mãe Daubechies 8. Na Figura 13 é possível observar o resultado após aplicação da função ECDF com o teste Mann-Whitney sobre os descritores aprendidos na camada de *flatten* da CNN e coeficientes wavelet das sub-bandas de detalhes. Nesses dados pode ser observado que 40% das amostras possuem um p -valor menor ou igual a 0,05, uma redução de 10% se comparado com as características apenas da ResNet-50. Esse valor mostra que a combinação das características de detalhe pode piorar a separabilidade dos dados.

Nas Tabelas 3 e 4 são apresentados os resultados da classificação da rede ResNet-

Tabela 3 – Avaliação da acurácia para a combinação de características da ResNet-50 e sub-bandas de detalhe da wavelet.

Classificadores	40X	100X	200X	400X
RF	77,0%	76,9%	85,7%	75,6%
SVM	88,6%	88,4%	89,7%	83,3%

Tabela 4 – Avaliação da AUC para a combinação de características da ResNet50 e sub-bandas de detalhe da Wavelet com a função *Daubechies 8*.

Classificadores	40X	100X	200X	400X
RF	0,647	0,652	0,769	0,642
SVM	0,830	0,831	0,842	0,762

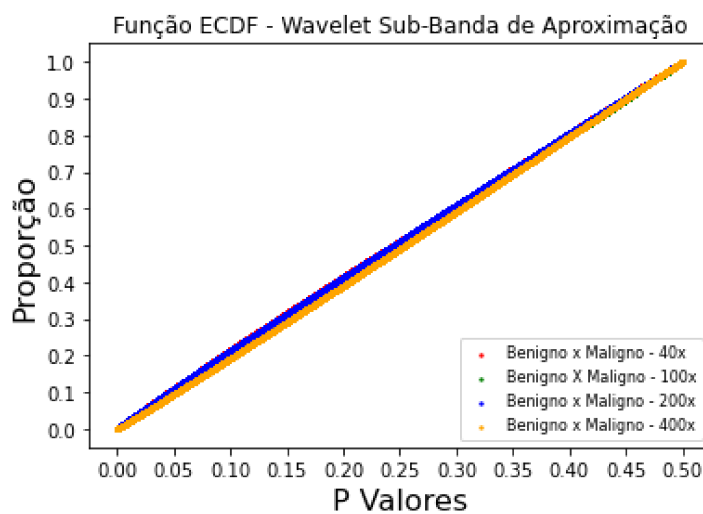


Figura 15 – ECDF para as características extraídas do modelo ResNet-50 combinadas com as sub-bandas de aproximação da transformada wavelet.

50 e descritores wavelet das sub-bandas de detalhe. Nesta avaliação, observa-se que os valores das medidas foram, no geral, inferiores aos resultados do *baseline* em todas as métricas para o classificador RF. No entanto, para o classificador SVM, não houve mudança significativa nas métricas de acurácia e AUC e em alguns níveis de magnificação. O modelo apresentou uma pequena melhora no desempenho, principalmente na métrica AUC em imagens com magnificação de 40x, onde o *baseline* atingiu uma área de 0,752 e com a combinação das características este valor subiu para 0,830.

Na Figura 15 é possível observar o resultado após aplicação da Função de Distribuição Empírica Acumulativa com o teste Mann-Whitney sobre os descritores aprendidos na camada de *flatten* da CNN e coeficientes wavelet da sub-banda de aproximação. Nesses dados pode ser observado que apenas 10% dos dados possuem *p*-valor menor ou igual a 0,05.

Nas Tabelas 5 e 6 são apresentados os resultados da classificação da rede ResNet-50 e descritores wavelet da sub-banda de aproximação. Nesta avaliação observa-se que os

Tabela 5 – Acurácia para a combinação de características da ResNet-50 e sub-banda de aproximação da transformada wavelet.

Classificadores	40X	100X	200X	400X
RF	77,0%	80,2%	86,8%	79,3%
SVM	88,6%	88,4%	89,5%	83,1%

Tabela 6 – Avaliação da AUC para a combinação de características da ResNet-50 e sub-banda de aproximação.

Classificadores	40X	100X	200X	400X
RF	0,684	0,707	0,792	0,703
SVM	0,827	0,830	0,839	0,759

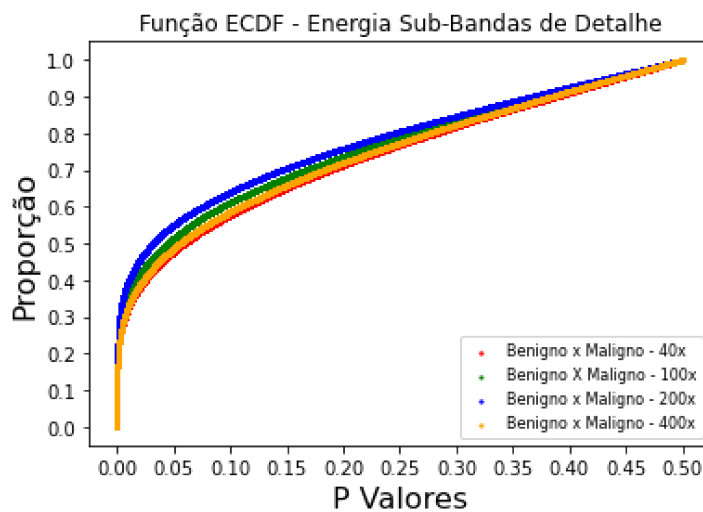


Figura 16 – ECDF para as características extraída do modelo ResNet50 combinadas com a energia extraída das sub-bandas de detalhe da aplicação da Transformada Wavelet Daubechies 8.

valores das medidas foram inferiores aos resultados do *baseline*. No entanto, é possível notar uma melhora nos resultados da acurácia e AUC do classificador RF em relação a combinação das características da ResNet-50 com as sub-bandas de detalhe.

A Função de Distribuição Empírica Acumulativa com o teste Mann-Whitney sobre os descritores aprendidos na etapa de *flatten* da CNN e descritores de energia *wavelet* das sub-bandas de detalhe é apresentada na Figura 16. Nesses dados pode ser observado que cerca de 50% das amostras possuem um *p*-valor menor ou igual a 0,05, uma distribuição semelhante a combinação de características da ResNet-50 com a energia extraída das sub-bandas de detalhe. Para o grupo de benigno versus maligno com 200x a separação é maior em relação as outras comparações.

Nas Tabelas 7 e 8 são apresentados os resultados da classificação da rede ResNet-50 e descritores wavelet da sub-banda de aproximação. Nesta avaliação observa-se que os valores das medidas foram estatisticamente idênticos aos obtidos no *baseline*, com

Tabela 7 – Avaliação da acurácia para a combinação de *features* da ResNet-50 e energia das sub-bandas de detalhe da transformada wavelet Daubechies 8.

Classificadores	40x	100x	200x	400x
RF	80,6%	81,0%	87,5%	80,4%
SVM	88,4%	88,4%	89,5%	83,1%

Tabela 8 – Avaliação da AUC para a combinação de *features* da ResNet-50 e energia das sub-bandas de detalhe da transformada wavelet Daubechies 8.

Classificadores	40x	100x	200x	400x
RF	0,710	0,717	0,798	0,713
SVM	0,827	0,830	0,839	0,759

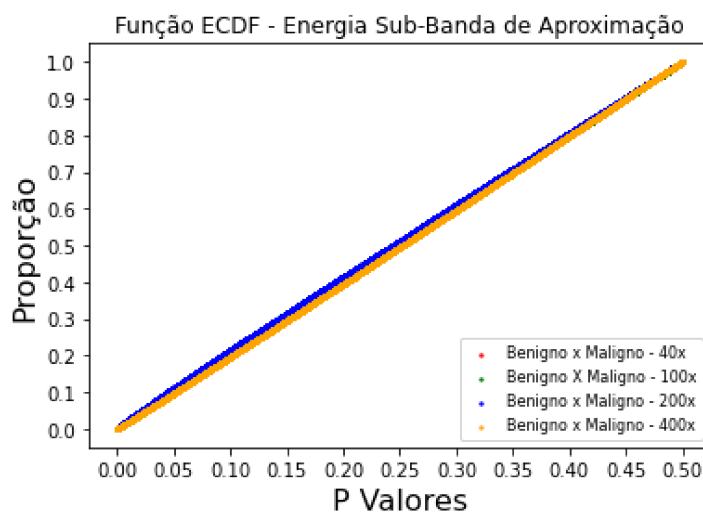


Figura 17 – ECDF para as características extraída do modelo ResNet-50 combinadas com a energia extraída da sub-banda de aproximação da aplicação da transformada wavelet Daubechies 8.

exceção da AUC nas imagens com magnificação de 40X para o classificador SVM. Essas pequenas alterações podem ser originárias da natureza não-determinística dos algoritmos de classificação empregados.

Com os descritores aprendidos na etapa de *flatten* da CNN e descritores de energia wavelet da sub-banda de aproximação, na Figura 17, tem-se o resultado após aplicação da Função de Distribuição Empírica Acumulativa com o teste Mann-Whitney. Esses dados mostram que apenas 10% dos dados possuem *p*-valor de menor ou igual a 0,05.

Nas Tabelas 9 e 10 são apresentados os resultados da classificação da rede ResNet-50 e descritores wavelet da sub-banda de aproximação para informações de energia. Nesta avaliação observa-se que os valores das medidas foram estatisticamente idênticos aos obtidos no *baseline*, com exceção da AUC nas imagens com magnificação de 40x para o classificador SVM. Os resultados foram muito semelhantes aos obtidos na classificação utilizando a combinação de características da ResNet-50 com a energia calculada sobre

Tabela 9 – Avaliação da acurácia para a combinação de *features* da ResNet-50 e energia da sub-banda de aproximação da transformada wavelet Daubechies 8.

Classificadores	40x	100x	200x	400x
Random Forest	80,9%	81,0%	87,2%	80,4%
SVM	88,4%	88,4%	89,5%	83,1%

Tabela 10 – Avaliação da AUC para a combinação de *features* da ResNet-50 e energia da sub-banda de aproximação da transformada wavelet Daubechies 8.

Classificadores	40X	100X	200X	400X
Random Forest	0,710	0,718	0,799	0,718
SVM	0,827	0,830	0,839	0,759

as sub-bandas de detalhe.

4.3 Considerações Finais

As CNNs têm sido muito utilizadas na literatura para a classificação de imagens histológicas, porque não há a necessidade de criar métodos de extração *handcrafted* que precisem ser adaptados para cada problema. No entanto, alguns métodos *handcrafted* foram testados na literatura por sua capacidade de extrair informações de textura de imagens histológicas, sendo um dos principais a transformada wavelet. Para tanto, este trabalho analisou a combinação dos descritores *handcrafted* e não-*handcrafted* para a classificação de imagens histológicas de câncer de mama.

Quando se analisa a acurácia dos resultados, nota-se que a SVM obtiveram os melhores resultados em todos os níveis de magnificação, com uma média de 88,5% para as imagens com 40X de magnificação, 88,4% para as de 100X, de 89,5% para as de 200X e 83,1% para as 400X. Ainda observa-se que não houve uma variação para o uso dos descritores não-*handcrafted*, ou combinados com os descritores *handcrafted* extraídos com a aplicação da transformada wavelet. Já o *ensemble* com o modelo RF como classificador, obteve resultados abaixo do que os alcançados pela SVM. Nota-se que o modelo forneceu valores mais baixos de acurácia quando a dimensionalidade do conjunto de dados foi elevada. Isso pode ser observado com os descritores combinados de detalhe da wavelet e os extraídos da ResNet-50. Em contra partida, a ResNet-50 obteve resultados mais homogêneos para todos os níveis de magnificação, com uma pequena perda de acurácia nas imagens de 400x.

Em relação a AUC, o melhor resultado foi atingido no *baseline* com o *softmax* da ResNet-50, com uma área média de 0,908, valor máximo de 0,936 para as imagens com 200x de magnificação e mínimo de 0,870 para as com 400x de magnificação. Todos os outros modelos, apresentaram taxas menores ao obtido com a ResNet-50. Um valor

médio de 0,7213 para o modelo RF e 0,801 para o SVM. Como o conjunto de dados é desbalanceado e nenhuma técnica de aumento de dados foi aplicada, a AUC pode ser considerada uma métrica mais significativa para o desempenho geral do modelo.

5 Conclusão

Para este trabalho foi idealizado um modelo de classificação que combinasse descritores *handcrafted* e *non-handcrafted* para a classificação de imagens histológicas de tumores de mama. O objetivo era investigar se a utilização de diferentes tipos de descritores combinados, melhoraria o desempenho da classificação. No entanto, durante a análise dos resultados, avaliou-se que modelo proposto não resultou em melhorias de desempenho em relação ao *baseline*. E no classificador RF, o aumento da dimensionalidade gerado pela concatenação dos coeficientes de detalhe, diminuiu de 81% para 77% a acurácia da classificação.

Verificou-se também que a utilização dos descritores *handcrafted* extraídos pela aplicação da transformada wavelet não contribuiu para o desempenho da classificação. Ainda em relação à transformada wavelet, analisou-se que a utilização dos coeficientes de aproximação diminuiu o *p*-valor das características, diminuindo a quantidade dos vetores de *features* que são utilizados para discriminar os grupos de lesões, o que explica a queda de desempenho da classificação. No comparativo entre utilizar os vetores de subbandas inteiros ou aplicar a função de energia, concluiu-se que não houveram diferenças entre a sua utilização e que o fator determinante para o desempenho da classificação foi o vetor de características *no-handcrafted*.

Em relação aos níveis de magnificação das imagens, conforme foram analisados e testados de maneira independente, observou-se que a magnificação de 200x apresentou as melhores taxas gerais de acurácia e área AUC, com 89,5% (utilizando o SVM) e área 0,936 (utilizando a *softmax*).

5.1 Desafios Encontrados

Durante o desenvolvimento do trabalho, diversos fatores contribuíram para que se tornasse uma tarefa desafiadora. Em especial, uma pandemia global que dificultou a comunicação e mudou pra sempre as formas de trabalho. O desafio de aprender sobre tópicos fora da área de domínio da Ciência da Computação, como histologia e como as lesões de tumores de mama se apresentam em imagens digitais, aprender sobre coloração e sobre as etapas do processamento digital de imagens. Além disso, encontrar uma arquitetura de Rede Neural Convolutiva adequada, ajustar os hiper-parâmetros para produzir bons vetores de características *non-handcrafted* consumiu muito tempo e a falta de um equipamento com GPU, dificultou a análise de modelos que demandavam de muito tempo e esforço computacional para serem treinados.

No entanto, todos os desafios encontrados possibilitaram crescimento pessoal e profissional, e espera-se que os conhecimentos adquiridos no desenvolvimento sejam úteis para minha carreira.

5.2 Trabalhos Futuros

Os descritores *handcrafted* analisados neste trabalho não contribuíram para a melhoria do desempenho de classificação das imagens histológicas de tumores de mama. Como é inegável a existência de diferenças morfológicas no tecido, é necessário um conjunto de dados maior, para que modelos especializados possam ser treinados para este domínio específico de imagens. Além disso, diminuir a dimensionalidade das características *no-handcrafted* pode ser um bom ponto de partida para analisar as combinações com descritores *handcrafted*. Espera-se também, que este trabalho contribua para o desenvolvimento de sistemas CAD, para melhorar o diagnóstico preciso e precoce das lesões de câncer de mama.

Referências

- ABADI, M. et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>. Citado na página 27.
- AL, L. et. Handwritten digit recognition: Applications of neural network chips and automatic learning. *IEEE Communications Magazine, IEEE*, v., v. 27, 1989. Disponível em: <<https://doi.org/10.1109/35.41400>>. Citado 3 vezes nas páginas 3, 19 e 20.
- ARAÚJO, F. H. D. Convnets na caracterização, recuperação e ranqueamento de células. *Universidade Federal do Ceará*, 2018. Citado na página 20.
- BAYRAMOGLU, N.; KANNALA, J.; HEIKKILÄ, J. Deep learning for magnification independent breast cancer histopathology image classification. In: *2016 23rd International Conference on Pattern Recognition (ICPR)*. [S.l.: s.n.], 2016. p. 2440–2445. Citado na página 21.
- CANCER, A. *Breast Anatomy*. 2021. Disponível em: <http://www.aboutcancer.com/breast_ca_org_anatomy.jpg>. Citado 2 vezes nas páginas 3 e 11.
- CAUDILL, M. Neural nets primer, part vi. *AI Expert*, v. 4, n. 2, p. 61–67, 1989. Citado na página 18.
- CONCI, A.; AZEVEDO, E.; LETA, F. R. *Computação Gráfica*. [S.l.: s.n.], 2021. ISBN 9788535223293. Citado 2 vezes nas páginas 14 e 16.
- DENG, J. et al. Imagenet: A large-scale hierarchical image database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2009. p. 248–255. Citado na página 27.
- DIMITROPOULOS, K. et al. Grading of invasive breast carcinoma through grassmannian vlad encoding. *PLOS ONE*, Public Library of Science, v. 12, n. 9, p. 1–18, 09 2017. Disponível em: <<https://doi.org/10.1371/journal.pone.0185110>>. Citado na página 21.
- DUARTE, Y. A. S.; NASCIMENTO, M. Z. do. *Algoritmos baseados em aproximações LBP em domínio Wavelet aplicados em mamogramas*. 2014. Disponível em: <<https://bit.ly/3hc3kXE>>. Citado na página 8.
- EVGENIOU, T.; PONTIL, M. Support vector machines: Theory and applications. In: . [S.l.: s.n.], 2001. v. 2049, p. 249–257. Citado na página 29.
- FUJIEDA, S.; TAKAYAMA, K.; HACHISUKA, T. Wavelet convolutional neural networks for texture classification. *CoRR*, abs/1707.07394, 2017. Disponível em: <<http://arxiv.org/abs/1707.07394>>. Citado na página 8.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento digital de imagens*. [S.l.: s.n.], 2007. v. 3ª Edição. Citado 2 vezes nas páginas 14 e 15.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. <<http://www.deeplearningbook.org>>. Citado 2 vezes nas páginas 19 e 20.

- HAN, Z. et al. Breast cancer multi-classification from histopathological images with structured deep learning model. *Scientific Reports*, v. 7, n. 1, p. 4172, Jun 2017. ISSN 2045-2322. Disponível em: <<https://doi.org/10.1038/s41598-017-04075-z>>. Citado na página 22.
- HE, K. et al. Deep residual learning for image recognition. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. Citado 2 vezes nas páginas 20 e 26.
- INCA. *Câncer de mama*. 2021. Disponível em: <<https://www.inca.gov.br/tipos-de-cancer/cancer-de-mama>>. Citado 2 vezes nas páginas 7 e 10.
- JOYSEEREE, R. et al. Fusing learned representations from riesz filters and deep cnn for lung tissue classification. *Medical Image Analysis*, v. 56, p. 172 – 183, 2019. ISSN 1361-8415. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S136184151830255X>>. Citado na página 8.
- KAUSHAL, C. et al. Recent trends in computer assisted diagnosis (cad) system for breast cancer diagnosis using histopathological images. *IRBM*, v. 40, n. 4, p. 211–227, 2019. ISSN 1959-0318. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1959031818301969>>. Citado 2 vezes nas páginas 3 e 13.
- KINGMA, D. P.; BA, J. *Adam: A Method for Stochastic Optimization*. 2017. Citado na página 27.
- KO, B. C.; KIM, S. H.; NAM, J.-Y. X-ray image classification using random forests with local wavelet-based cs-local binary patterns. *Journal of digital imaging*, Springer-Verlag, v. 24, n. 6, p. 1141–1151, Dec 2011. ISSN 1618-727X. 21487837[pmid]. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/21487837>>. Citado na página 29.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F. et al. (Ed.). *Advances in Neural Information Processing Systems 25*. Curran Associates, Inc., 2012. p. 1097–1105. Disponível em: <<http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>>. Citado na página 20.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, v. 86, n. 11, p. 2278–2324, 1998. Citado na página 20.
- LEE, G. R. et al. *PyWavelets/pywt: PyWavelets v1.0.3*. Zenodo, 2019. Disponível em: <<https://doi.org/10.5281/zenodo.2634243>>. Citado na página 27.
- MISITI, M. et al. Matlab wavelet toolbox - user's guide. p. 9–11, 2008. Citado na página 16.
- MOREIRA JUNIOR, W. L.; NASCIMENTO, M. Z. do. *Uma nova abordagem de descritor de textura baseada em transformada ripple para classificação de lesões de mama*. 2018. Disponível em: <http://biblioteca.ufabc.edu.br/index.php?codigo_sophia=110468>. Citado 3 vezes nas páginas 3, 7 e 17.

- NASCIMENTO, M. Z. do et al. Classification of masses in mammographic image using wavelet domain features and polynomial classifier. *Expert Systems with Applications*, v. 40, n. 15, p. 6213–6221, 2013. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417413002844>>. Citado na página 27.
- OLIVEIRA, H. de. *Análise de Sinais para Engenheiros: Uma abordagem via WAVELETS*. [S.l.: s.n.], 2007. ISBN 978-85-7452-283-8. Citado na página 16.
- PEDREGOSA, F. et al. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, v. 12, n. 85, p. 2825–2830, 2011. Disponível em: <<http://jmlr.org/papers/v12/pedregosa11a.html>>. Citado 2 vezes nas páginas 29 e 30.
- RAMOS, R. P.; NASCIMENTO, M. Z. do; PEREIRA, D. C. Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms. *Expert Systems with Applications*, v. 39, n. 12, p. 11036–11047, 2012. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417412004903>>. Citado na página 29.
- SEUÁNEZ, H. N. Introdução à rede nacional de câncer familiar. *Rede nacional de câncer familiar: manual operacional.*, v. 2, p. 9–14, Jul 2009. Disponível em: <http://www.inca.gov.br/inca/Arquivos/publicacoes/Cancer_Familiar_fm.pdf>. Citado na página 10.
- SIMONYAN, K.; ZISSERMAN, A. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 2014. Citado 2 vezes nas páginas 3 e 28.
- SPANHOL, F. A. et al. A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, v. 63, n. 7, p. 1455–1462, July 2016. ISSN 0018-9294. Citado 7 vezes nas páginas 3, 7, 11, 15, 21, 23 e 24.
- SZEGEDY, C. et al. Going deeper with convolutions. In: *Computer Vision and Pattern Recognition (CVPR)*. [s.n.], 2015. Disponível em: <<http://arxiv.org/abs/1409.4842>>. Citado na página 20.
- ZUIN, G. L.; VELOSO, A. A. Uso de adaptação de domínio e informação contextual em sistemas de pergunta-resposta. nov. 2017. Disponível em: <<https://repositorio.ufmg.br/handle/1843/JCES-AVPH4J>>. Citado 2 vezes nas páginas 3 e 19.