



Programa de Pós-Graduação em

Computação Aplicada

Mestrado/Doutorado Acadêmico

Pedro Clarindo da Silva Neto

BGNDL: Arquitetura de Deep Learning para diferenciação da
proteína Biglycan em tecido mamário com e sem câncer

São Leopoldo, 2022

Pedro Clarindo da Silva Neto

BGNDL: ARQUITETURA DE DEEP LEARNING PARA DIFERENCIAÇÃO DA
PROTEÍNA BIGLYCAN EM TECIDO MAMÁRIO COM E SEM CÂNCER

Dissertação apresentada como requisito parcial
para a obtenção do título de Mestre pelo
Programa de Pós-Graduação em Computação
Aplicada da Universidade do Vale do Rio dos
Sinos — UNISINOS

Orientador:
Prof. Dr. Rafael Kunst

Co-orientador:
Prof. Dr. Jorge Luis Victoria Barbosa

São Leopoldo
2022

S586b Silva Neto, Pedro Clarindo da.
BGNDL : arquitetura de deep learning para diferenciação da proteína biglycan em tecido mamário com e sem câncer / por Pedro Clarindo da Silva Neto. – 2022.
62 f. : il. ; 30 cm.

Dissertação (mestrado) — Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Computação Aplicada, São Leopoldo, RS, 2022.
Orientador: Dr. Rafael Kunst.
Co-orientador: Dr. Jorge Luis Victoria Barbosa.

1. Deep learning. 2. Câncer de mama. 3. Biomarcador.
4. Biglycan. I. Título.

CDU: 004.85:618.19-006

(Esta folha serve somente para guardar o lugar da verdadeira folha de aprovação, que é obtida após a defesa do trabalho. Este item é obrigatório, exceto no caso de TCCs.)

Ao meu avô, Pedro Clarindo da Silva (*in memoriam*)

Se eu vi mais longe, foi por estar sobre ombros de gigantes — SIR ISAAC NEWTON

AGRADECIMENTOS

A Deus, pela dádiva da vida.

Ao Prof. Dr. Rafael Kunst, pela orientação, competência, profissionalismo e dedicação tão importantes. Agradeço por todos os ensinamentos compartilhados de forma admirável, pela parceria e amizade. Ao meu coorientador, Prof. Dr. Jorge Luis Victoria Barbosa, por toda a ajuda durante a realização deste trabalho.

Aos membros da banca examinadora, Prof. Dr. Ricardo Francalacci Savaris, Prof. Dr. Sandro Jose Rigo e Prof. Dr. Cristiano André da Costa, que tão gentilmente aceitaram participar e colaborar com esta dissertação.

Aos professores e funcionários do PPGCA, pela dedicação, competência, apoio e todo conhecimento compartilhado, em especial do Prof. Dr. Rodrigo Righi que acreditou no Minter em Mato Grosso.

Ao amigo Jovani Dalzochio, por toda ajuda e cooperação nas atividades desenvolvidas durante o mestrado.

Às minhas tias Jandira e Cristina, por todas as lições de amor, dedicação, abnegação e compreensão. Sinto-me orgulhoso e privilegiado por ter mães tão especiais.

À minha amada noiva Thiessa, por todo amor, compreensão e apoio em tantos momentos difíceis desta caminhada. Obrigado por permanecer ao meu lado.

Por fim, a todos aqueles que contribuíram, direta ou indiretamente, para a realização desta dissertação, o meu sincero agradecimento.

“Ninguém abre um livro sem que aprenda alguma coisa”.
(Anônimo)

RESUMO

A Inteligência Artificial e o Aprendizado de Máquina tornaram-se aliados importantes na área da saúde. Nesse contexto, Deep Learning tem oferecido suporte a tarefas médicas críticas, incluindo diagnóstico, previsão de resultados e resposta ao tratamento. As imagens histológicas, foco deste trabalho, são provenientes dos tecidos do corpo humano. O diagnóstico de muitas doenças, principalmente as malignas, depende da avaliação de cortes histológicos. Dentro do cenário das avaliações diagnósticas por imagem, existem variações. Na literatura, demonstrou-se que apesar da consistência dos resultados em um mesmo avaliador, existe uma diferença entre diferentes avaliadores. Segundo a literatura, as diferenças na percepção visual e no treinamento clínico podem levar a inconsistências nas opiniões diagnósticas e prognósticas uma vez que a análise patológica é naturalmente subjetiva. A coloração de rotina dos tecidos para estudo microscópico nem sempre é suficiente. Nestes casos são utilizados como complementos os marcadores biológicos, os biomarcadores. Nesse sentido, o crescente interesse pela pesquisa de biomarcadores tem sido percebido devido ao aumento dos custos de pesquisa e do tempo necessário para desenvolver um novo composto. Para que esses biomarcadores sejam utilizados em pesquisas, é necessário que o mesmo passe por um processo de validação, onde ele precisa ser medido em um sistema de teste, onde uma das propriedades, a sensibilidade do biomarcador, será avaliada neste trabalho. Com a exposição deste cenário, este trabalho descreveu, através do Deep Learning, a criação da arquitetura CNN que irá verificar, a partir de imagens histológicas com o biomarcador Biglycan, se há diferença entre a expressão do Biglycan entre tecidos com e sem câncer de mama. A associação de Deep Learning e da expressão da proteína Biglycan, através da intensidade da coloração do DAB, utilizando a deconvolução das cores é nova e necessária para validação do biomarcador. Nesse sentido, as principais contribuições deste trabalho são: Criação de um conjunto de dados original de imagens histológicas com e sem câncer de mama que foram submetidas à técnica de imunohistoquímica para determinar a expressão da proteína Biglycan, Automação do modelo de deconvolução de cores para analisar apenas imagens com expressão DAB e Desenvolvimento de uma arquitetura CNN que possa determinar se há diferença entre a expressão de Biglycan entre tecidos com e sem câncer de mama. As imagens da histologia da mama foram classificadas por uma média percentual superior a 93%, indicando que há diferença entre a expressão do biomarcador Biglycan entre os tecidos com e sem câncer de mama.

Palavras-chave: Deep Learning. Câncer de Mama. Biomarcador. Biglycan.

ABSTRACT

Artificial Intelligence and Machine Learning have become important allies in healthcare. In this context, Deep Learning has provided support for critical medical tasks, including diagnosis, outcome prediction, and treatment response. Histological images, the focus of this work, come from the tissues of the human body. The diagnosis of many diseases, especially malignant diseases, depends on the evaluation of histological sections. Within the scenario of diagnostic imaging evaluations, variations exist. In the literature, it has been shown that despite the consistency of results in the same rater, there is a difference between different raters. According to the literature, differences in visual perception and clinical training can lead to inconsistencies in diagnostic and prognostic opinions since pathological analysis is naturally subjective. Routine staining of tissues for microscopic study is not always sufficient. In these cases biological markers, the biomarkers, are used as complements. In this regard, the growing interest in biomarker research has increased due to rising research costs and the time required to develop a new compound. For these biomarkers to be used in research, it is necessary that they go through a validation process, where they need to be measured in a test system, where one of the properties, the sensitivity of the biomarker, will be evaluated in this work. With the exposure of this scenario, this work promoted, through Deep Learning, the creation of the CNN architecture that will check, from histological images with the biomarker Biglycan, if there is a difference between the expression of Biglycan between tissues with and without breast cancer. The association of Deep Learning and Biglycan protein expression by DAB staining intensity using color deconvolution is new and necessary for biomarker validation. In this sense, the main contributions of this work are: Creation of an original dataset of histological images with and without breast cancer that were subjected to the immunohistochemistry technique to determine Biglycan protein expression, Automation of the color deconvolution model to analyze only images with DAB expression and Development of a CNN architecture that can determine whether there is a difference between Biglycan expression between tissues with and without breast cancer. The breast histology images were classified by an average percentage greater than 93%, indicating that there is a difference between Biglycan biomarker expression between tissues with and without breast cancer.

Keywords: Deep Learning. Breast Cancer. Biomarker. Biglycan.

LISTA DE FIGURAS

Figura 1:	Técnicas de obtenção de imagens	34
Figura 2:	Fluxo das atividades executadas	42
Figura 3:	Exemplo das imagens disponibilizadas pelo HCPA, sendo A - Saudável e B - Câncer	43
Figura 4:	Imagens após Color Deconvolution, sendo A - Saudável e B - Câncer	44
Figura 5:	Processo de execução de script do color deconvolution	44
Figura 6:	Dados de treinamento listados em sequência	45
Figura 7:	Dados de treinamento listados após embaralhamento aleatório	46
Figura 8:	Implementação do Modelo Proposto	47
Figura 9:	Exemplos de imagens dos datasets utilizados	49
Figura 10:	Acurácia e Perda do modelo - Imagens sem deconvolução de cores	50
Figura 11:	Exemplo de Validação-Cruzada K-Fold = 5	50
Figura 12:	Matriz confusão e métricas - modelo BGNDL	51
Figura 13:	Curva ROC e AUC do modelo BGNDL	52
Figura 14:	Comparação modelos BGNDL e ResNet50	52
Figura 15:	Imagens classificadas erradas pelo modelo BGNDL	53

LISTA DE TABELAS

Tabela 1:	Redes de Deep Learning e suas aplicações	30
Tabela 2:	Camadas de saída e funções de perda e suas finalidades	32
Tabela 3:	Trabalhos Relacionados	40
Tabela 4:	Dataset - Conjunto de imagens de treino e teste	43

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Motivação	21
1.2	Definição do Problema	22
1.3	Objetivos	24
1.4	Metodologia	24
1.5	Organização da dissertação	25
2	FUNDAMENTAÇÃO TEÓRICA	27
2.1	Deep Learning	27
2.1.1	Redes Neurais Convolucionais	30
2.2	Visão Computacional aplicada à Medicina	32
2.3	Deteccção de Câncer por Imagem	34
2.3.1	Patologia Digital e uso de Deep Learning	35
2.3.2	Biglycan	35
3	TRABALHOS RELACIONADOS	37
4	ARQUITETURA DE <i>DEEP LEARNING</i> PARA DETECÇÃO DE CÂNCER DE MAMA COM BGN	41
4.1	Arquitetura Proposta	41
4.2	<i>Dataset</i> e Pré-processamento	42
4.3	Modelo e treinamento	45
4.4	Métricas de Avaliação	47
5	RESULTADOS	49
6	CONCLUSÃO	55
	REFERÊNCIAS	57

1 INTRODUÇÃO

Neste Capítulo são apresentadas motivações para a elaboração da dissertação (Seção 1.1), o problema de pesquisa (Seção 1.2), os objetivos geral e específicos (Seção 1.3) e, por fim, a metodologia do trabalho (Seção 1.4).

1.1 Motivação

Conforme Srinivasan e Srinivasna (2020), cerca de 400 milhões de pessoas não têm acesso aos serviços de saúde necessários, com uma média de um médico para 2.000 pessoas em países em desenvolvimento. Os autores apontam técnicas como Inteligência Artificial (IA) e Aprendizado de Máquina, do inglês, *Machine Learning* (ML) como aliadas no combate a doenças como câncer de pele, onde, só nos Estados Unidos, são diagnosticadas mais de 9500 pessoas por ano. De acordo com um relatório da Organização Mundial da Saúde (OMS), o câncer de mama é responsável por 22,9% dos cânceres diagnosticados e 13,7% das mortes relacionadas ao câncer em todo o mundo (LI et al., 2018).

De acordo com Lee et al. (2012), cerca de 1,38 milhão de mulheres são diagnosticadas com câncer de mama em todo mundo e 458.503 morrem da doença todos os anos. A incidência está aumentando em países de baixa e média renda, uma tendência que tem sido atribuída à melhora na expectativa de vida, urbanização e adoção de estilos de vida ocidentais. O câncer de mama é o câncer mais comum em mulheres em todo o mundo e 70% das mortes por câncer de mama ocorrem em mulheres de países de baixa e média renda. A América Latina tem cerca de 115 mil novos casos da doença a cada ano, com cerca de 50 mil surgindo no Brasil.

Chang et al. (2017) apontaram que em 2012 houve 1,7 milhão de casos de câncer de mama, e que é o câncer feminino mais comum em 140 de 184 países. Informam também que a detecção precoce do câncer de mama é um fator importante na sobrevivência e que a detecção de tem sido determinada pelo diagnóstico patológico de especialistas, que é influenciado pela experiência do médico e outros fatores externos. Relatam ainda que, para resolver este problema, métodos de análise assistida por computador têm sido aplicados em imagens médicas, incluindo algoritmos de ML, especialmente, a rede neural profunda.

Habuza et al. (2021) relata que a IA mostrou resultados promissores na detecção de doenças como hemorragia, acidente vascular cerebral ou câncer em seus estágios iniciais com mais precisão e com uma menor taxa de falsos positivos. De acordo com a *American Cancer Society*, a taxa de mortalidade por câncer nos Estados Unidos diminuiu continuamente nos últimos anos, conforme publicado em relatórios de estatísticas anuais.

Segundo Maier et al. (2019), Aprendizado Profundo, do inglês *Deep Learning* (DL), oferece um grande conjunto de novas ferramentas que são aplicáveis a muitos problemas no mundo do processamento de imagens médicas e cita que na conferência internacional de Computação de Imagens Médicas e Intervenção Assistida por Computador (MICCAI) em 2018, aproximada-

mente 70% de todas as publicações aceitas estavam relacionadas ao tópico de DL.

Para Fatima e Pasha (2017), ML é importante no diagnóstico auxiliado por computador e que o reconhecimento de padrões envolve fundamentalmente aprender com os exemplos. Citam que no campo da biomedicina, o reconhecimento de padrões e o ML prometem maior precisão na percepção e diagnóstico de doenças e que também promovem a objetividade do processo de tomada de decisão.

De acordo com Srinidhi, Ciga e Martel (2021), em 2011, pesquisadores demonstraram que características extraídas de imagens histológicas podem auxiliar na descoberta de novos aspectos biológicos do tecido canceroso. Em seu artigo, os autores ainda citam que as pesquisas mostraram que recursos extraídos de imagens de patologia digital são complementares aos dados genômicos. Levando ao cenário onde esses avanços geraram um crescente interesse no uso de biomarcadores extraídos de imagens de patologia digital para medicina de precisão, particularmente em oncologia.

Esteva et al. (2021) cita que nos últimos anos, o número de publicações que aplicam técnicas de visão computacional a imagens médicas estáticas cresceu de centenas para milhares, onde algumas áreas recebem atenção substancial - radiologia, patologia, oftalmologia e dermatologia - devido à natureza de reconhecimento de padrão visual das tarefas de diagnóstico nessas especialidades e à crescente disponibilidade de imagens altamente estruturadas.

Diante do exposto, foi possível observar que o câncer é uma doença que mata muitas pessoas anualmente no mundo todo e que o câncer de mama é responsável por uma quantidade significativa de casos. Também foi observado que a detecção precoce da doença é fator importante para sobrevivência. Técnicas assistidas por computador tem demonstrado serem aliadas nesses casos e o uso de ML vem sendo usado cada vez mais como ferramenta nesse processo. Com a exposição deste cenário, esta dissertação promoveu, com o uso de Deep Learning, a criação de uma arquitetura de CNN que irá verificar, a partir de imagens histológicas com o biomarcador Biglycan, se existe diferença entre a expressão do Biglycan entre tecidos com e sem câncer de mama e, posteriormente, irá verificar a intensidade deste biomarcador para alimentar os modelos de ML.

1.2 Definição do Problema

Existem variações entre as avaliações do diagnóstico por imagem. O estudo de Fuhrich, Lessey e Savaris (2013) demonstrou que apesar de haver coerência de resultados num mesmo avaliador, existe diferença entre avaliadores diferentes. Diferenças na percepção visual e no treinamento clínico podem levar a inconsistências nas opiniões diagnósticas e prognósticas (KIM et al., 2017). A análise patológica - baseada na inspeção visual de amostras de tecido ao microscópio - é inerentemente subjetiva por natureza (ESTEVA et al., 2021).

Como informado por Srinivasan e Srinivasna (2020), IA e ML tem se tornando importantes aliados na área da saúde. Os autores indicam que algoritmos estão sendo treinados para re-

conhecer certas doenças e distúrbios quase que instantaneamente, usando redes neurais, como Redes Neurais Convolucionais, do inglês *Convolutional Neural Networks* (CNN). Neste cenário, DL pode apoiar tarefas médicas críticas, incluindo diagnóstico, prognóstico de resultados e resposta ao tratamento, segmentação de patologia, monitoramento de doenças e assim por diante (ESTEVA et al., 2021).

As imagens histológicas, foco desta pesquisa, tem sua origem nos tecidos do corpo humano. O diagnóstico de muitas doenças, especialmente as malignas, depende da avaliação de cortes histológicos (KATHER et al., 2015). A coloração de rotina dos tecidos para estudo em microscópio nem sempre são suficientes, nestes casos também são utilizados marcadores biológicos, os biomarcadores, que possuem características que são medidas como um indicador de um processo biológico ou patológico, ou de uma resposta farmacológica (HUNTER et al., 2010).

Somado a isso, o crescente interesse na pesquisa de biomarcadores considera o aumento dos custos de pesquisa e desenvolvimento e que o tempo e o custo necessários para desenvolver um novo composto aumentaram ainda mais nos últimos anos. Contudo, para que um biomarcador seja válido, é necessário que seja medido em um sistema de teste analítico com características de desempenho bem estabelecidas, onde uma das propriedades, a sensibilidade do biomarcador - capacidade de ser medido com precisão adequada e com magnitude de mudança suficiente, para torná-lo sensível o suficiente para refletir uma mudança significativa em desfechos clínicos importantes (HUNTER et al., 2010) - será avaliada neste trabalho.

Atualmente, o Biglycan é apontado como um potencial biomarcador para doenças (ZHAO et al., 2020), os autores descrevem no texto, que a sua distribuição não é uniforme nos órgão e tecidos, mas se localiza em tipos de células especializadas e que os padrões da expressão da Biglycan *in vivo* são alterados de acordo com as doenças. A expressão do RNA mensageiro da Biglycan foi avaliada em *in vitro* (BISCHOF et al., 2013) e (SHAO et al., 2021), em modelos animais (CONG et al., 2021) e em *in silico* (ZHAO et al., 2020).

Tendo vista o relato de Esteva et al. (2017), onde DL foi capaz de aprender com 129.450 imagens de 2.032 doenças e diagnosticar no mesmo nível que 21 dermatologistas certificados, e o de Srinivasan e Srinivasna (2020) que informou que existe, atualmente, uma alta demanda para atendimentos em serviços de saúde e que não conseguem ser supridos pela falta de profissionais, observa-se que o uso de ML, como ferramenta para medicina um forte aliado.

Com base nisso, abre-se a seguinte questão de pesquisa:

Como seria um método, através de imagens histológicas com o biomarcador Biglycan, utilizar *Machine Learning*, através de Redes Neurais Convolucionais, para determinar se a expressão desse biomarcador é diferente entre tecidos mamários com e sem malignidade?

Esta questão pode ser respondida ao desenvolver uma arquitetura, utilizando Redes Neurais Convolucionais, a partir de um *dataset* original criado com essas imagens e avaliando os resulta-

dos dessa rede. Vários estudos são apresentados, apontando as CNNs como uma estrutura mais adequada para executar essa tarefa, bem como a utilização do Biglycan como biomarcador.

1.3 Objetivos

Este trabalho propõe uma arquitetura de ML, que, através de DL e o algoritmo de CNN, verificar, dentro do contexto do câncer de mama, a relação da proteína Biglycan e o câncer de mama como um potencial biomarcador para esta doença. Para atingir este objetivo, este trabalho incidirá sobre os seguintes objetivos específicos:

- Compreender o estado-da-arte de ML;
- Propor a arquitetura da CNN;
- Construir um *dataset* original, com imagens coletadas com base em parceria com instituição de saúde;
- Construir um modelo para implementar a arquitetura proposta;
- Realizar testes com o modelo implementado;
- Avaliar a arquitetura.
- Verificar se existe diferença entre a expressão do Biglycan entre tecidos com e sem câncer de mama.

1.4 Metodologia

Esta pesquisa, do ponto de vista de sua natureza, objetiva gerar conhecimentos para aplicação prática dirigidos à solução de problemas específicos e classifica-se como aplicada (SILVA; MENEZES, 2001). A pesquisa é classificada como exploratória de acordo com a classificação dos objetivos. Estas pesquisas têm como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses (GIL, 2002). Quanto à abordagem do problema, a pesquisa classifica-se como quantitativa, pois considera que tudo pode ser quantificável, o que significa traduzir em números opiniões e informações para classificá-las e analisá-las, neste caso, as métricas para que foram aplicadas para analisar e avaliar os resultados do estudo (SILVA; MENEZES, 2001).

O primeiro passo para atingir os objetivos propostos é fazer uma revisão bibliográfica da temas relacionados com DL, Câncer de Mama e biomarcadores. Este estudo servirá de suporte para entender como funciona o que vem sendo já desenvolvido dentro da área de ML, para identificar quais técnicas são empregadas e entender seu uso. A segunda etapa será propor uma arquitetura, explicando os elementos que a compõem como eles ajudam a atingir o objetivo de

forma satisfatória. Uma terceira etapa é desenvolver um modelo da arquitetura com a funcionalidade mínima para validar a operação por meio de um *dataset* criado a partir de imagens fornecidas pelo serviço de patologia do Hospital de Clínicas de Porto Alegre. E, por último, uma quarta etapa para avaliar o desempenho da arquitetura.

1.5 Organização da dissertação

Neste capítulo, foram apresentadas a Motivação, Definição de Problemas, Objetivos e Metodologia. No Capítulo 2 é descrita a fundamentação teórica utilizada para o desenvolvimento da dissertação. São apresentados os tópicos acerca de Deep Learning, Visão Computacional aplicada à Medicina e a Detecção do Câncer por Imagem.

O capítulo 3 é dedicado a apresentar os trabalhos relacionados com o tema da dissertação. O Capítulo 4 apresenta a arquitetura proposta para a CNN que irá realizar a detecção do câncer, a etapa de pré-processamento, criação de dataset, criação do modelo, treinamento e métricas. No capítulo 5 são apresentados os resultados obtidos pelo treinamento do modelo e, por fim, no Capítulo 6 são apresentadas as conclusões e os futuros trabalhos.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta a fundamentação teórica acerca do tema principal de pesquisa desta dissertação. Com isso, são apresentados os conceitos e o histórico de DL na Seção 2.1, conceitos e definições de visão computacional aplicada à medicina na Seção 2.2 e uma introdução sobre o câncer e a detecção por imagem na Seção 2.3.

2.1 Deep Learning

De acordo Messaoud et al. (2020) ML é uma área de estudo que se encontra na reunião de vários campos, como IA, estatística, ciência cognitiva, teoria da probabilidade, otimização, sinal e informação. Os sistemas de ML são usados para identificar objetos em imagens, transcrever fala em texto, combinar notícias, publicações ou produtos com os interesses dos usuários e selecionar resultados de pesquisa relevantes (LECUN; BENGIO; HINTON, 2015). De acordo com T.K., Annavarapu e Bablani (2021) e Messaoud et al. (2020), há quatro abordagens ou categorias quando se trata de ML: Aprendizagem supervisionada, Aprendizagem semi-supervisionada, Aprendizagem não supervisionada e Aprendizagem por reforço.

O aprendizado supervisionado é útil nos casos em que uma propriedade de rótulo está disponível para um conjunto de dados individual. A aprendizagem semi-supervisionada possui poucos dados rotulados e poucas saídas conhecidas. A Aprendizagem não supervisionada é utilizada nos casos em que o desafio é identificar relacionamentos implícitos em um determinado conjunto de dados não rotulado. E, por último, a Aprendizagem por reforço, se concentra em experiências baseadas na tomada de decisão sequencial, usando recompensas onde o feedback são ações (T.K.; ANNAVARAPU; BABLANI, 2021).

Segundo Messaoud et al. (2020), para desenvolver um modelo preditivo, o aprendizado de máquina supervisionado utiliza técnicas ou modelos de classificação e regressão. As técnicas de classificação, geralmente são utilizadas para prever uma resposta onde os dados podem ser categorizados, rotulados ou separados em grupos ou classes específicas. Por exemplo, os aplicativos de reconhecimento de manuscrito para reconhecer letras e números. As técnicas de regressão, são usadas para prever respostas contínuas, o que significa trabalhar com uma faixa de dados ou se a natureza da resposta for um número real, como a temperatura. Algumas das técnicas utilizadas na tarefa de classificação são Árvores de Decisão, Redes Neurais, Regras de Classificação, Aprendizado Bayesiano, Máquinas de Vetores e Aprendizado Tardio (SANTOS, 2016).

Conforme Santos (2016), as Redes Neurais tiveram sua origem a partir da analogia de neurônios biológicos, são modelos de computação para o processamento de informações e são particularmente úteis para a identificação da relação fundamental entre um conjunto de variáveis ou padrões nos dados. O princípio da Rede Neural é criar um raciocínio baseado no que ele aprendeu no passado, utilizando dados de treinamento apresentados em pares de entrada-alvo

que são apresentados à rede um após o outro durante a fase de aprendizado, a fim de aprender. As entradas podem ser vistas como "perguntas" e os alvos como "respostas", sendo que cada vez que um par de pergunta e resposta é apresentado à Rede Neural, esta sabe a resposta para uma determinada questão (MESSAOUD et al., 2020).

Para LeCun, Bengio e Hinton (2015), as técnicas convencionais de ML apresentam limitações em sua capacidade de processar dados naturais em sua forma bruta. Ainda de acordo com o autor, cada vez mais tem-se utilizado uma classe de técnicas chamada DL, que se trata de um ramo de ML (DONG; WANG; ABBAS, 2021) e (KAJAREE DAS, 2017). O movimento sobre as redes neurais teve seu início na década de 1940, quando Warren McCulloch e Walter Pitts, escreveram um artigo sobre como os neurônios podem funcionar, onde propuseram uma rede neural básica com circuitos elétricos.

Na mesma década, Donald Hebb teorizou que as vias neurais são fortalecidas cada vez que são usadas. Na década de 1950, Nathaniel Rochester pesquisou uma rede neural abstrata simulada em computadores IBM 704. Em 1956, John McCarthy, Marvin L. Minsky, Nathaniel Rochester e Claude E. Shannon trabalharam juntos no projeto *Dartmouth Summer Research project on Artificial Intelligence*, onde forneceram um salto perspicaz na pesquisa de IA. Em 1957, John Von Neumann propôs que relés telegráficos ou tubos a vácuo podem ser usados para imitar a função simples dos neurônios. Em 1958, Frank Rosenblatt começou a trabalhar no Perceptron, que é a rede neural mais estabelecida ainda em uso hoje (DONG; WANG; ABBAS, 2021).

Seguidas pesquisas deram continuidade no avanço e no desenvolvimento das redes neurais. Em 1959, Bernard Widrow e Marcian Hoff de Stanford criaram os modelos ADALINE e MADALINE. MADALINE foi a rede neural primária a ser aplicada a um problema do mundo real, sendo um canal adaptável que eliminava ecos nas linhas telefônicas. Esses fatos causaram grande euforia das aplicações das redes neurais, contudo, haviam muitas restrições de hardware, e toda essa euforia acabou causando impactos negativos às pesquisas. O interesse pela pesquisa na área volta na década de 1980, quando John Hopfield apresentou um trabalho onde a metodologia era fazer máquinas mais valiosas, utilizando linhas bidirecionais. Antes, as associações entre neurônios eram apenas uma única via.

Em 1982, ocorre uma importante Conferência conjunta EUA-Japão sobre Redes Neurais Cooperativas/Competitivas. Em 1985, o *American Institute of Physics* deu início ao encontro anual - Redes Neurais para Computação. Em 1987, a primeira Conferência Internacional sobre Redes Neurais do *Institute of Electrical and Electronic Engineers* (IEEE). Em 1997, uma estrutura de rede neural recorrente, *Long Short-Term Memory* (LSTM) foi proposta por Schmidhuber e Hochreiter. Este modelo além de circular focos de informações individuais (por exemplo, imagens), mas também uma sequência inteira de dados (por exemplo, fala ou vídeo). Em 1998, Yann LeCun publicou a Aprendizagem Baseada em Gradiente Aplicada ao Reconhecimento de Documentos, que foi um passo importante no aprendizado de dados (DONG; WANG; ABBAS, 2021).

De acordo com Litjens et al. (2017), até o final da década de 1990, as técnicas supervisionadas, se tornaram cada vez mais populares na análise de imagens médicas. Esse reconhecimento de padrões ou abordagem de ML ainda é muito popular e forma a base de muitos sistemas de análise de imagens médicas comercialmente disponíveis. Os autores ainda informam que uma etapa crucial no projeto dos sistemas de imagens médicas é a extração de características discriminantes das imagens e que esse processo ainda é feito por pesquisadores humanos. Para os autores, a próxima etapa é permitir que os computadores aprendam os recursos que representam de maneira ideal os dados para o problema em questão. Este conceito está na base de muitos algoritmos de DL: modelos compostos por muitas camadas que transformam dados de entrada (por exemplo, imagens) em resultados (por exemplo, doença presente / ausente) enquanto aprende recursos de nível cada vez mais alto.

DL é uma das técnicas de ML que permite que modelos computacionais compostos de várias camadas de processamento aprendam representações de dados com vários níveis de abstração. Esses métodos melhoraram drasticamente o estado da arte em reconhecimento de fala, reconhecimento de objeto visual, detecção de objeto e muitos outros domínios (DONG; WANG; ABBAS, 2021) e (KAJAREE DAS, 2017).

Conforme Pouyanfar et al. (2018), tradicionalmente, a confiança na eficiência dos algoritmos de ML dependia da representação dos dados de entrada. Uma representação de dados ruim geralmente levava a um desempenho inferior em comparação com uma boa representação de dados. Em comparação, os algoritmos de DL realizam extração de recursos de forma automatizada, o que permite extrair características distintas com mínimo de conhecimento de domínio e esforço humano.

DL é uma classe de métodos de ML com vários níveis de representação. É composto por vários módulos simples, mas não lineares que cada um transforma a representação dos níveis anteriores em uma representação em um nível mais alto e ligeiramente mais abstrato. Com a composição de tais transformações, características e inferências muito complexas podem ser aprendido (HOSSEINI et al., 2020).

Sobre as classificações de DL, Hosseini et al. (2020) informa que existe três categorias principais de redes neurais: Redes Neurais Convolucionais (*Convolutional Neural Networks*), Redes Não-supervisionadas Pré-treinadas (*Pretrained Unsupervised Networks*) e Redes Neurais Recorrentes / Recursivas (*Recurrent/Recursive Neural Networks*). Já no estudo de Pouyanfar et al. (2018) são apresentadas algumas redes de DL, como: *Recursive Neural Network* (RvNN), *Recurrent Neural Network* (RNN), CNN e *Deep Generative Networks*.

Segundo Shrestha e Mahmood (2019), rede neural profunda consiste em várias camadas de nós. Diferentes arquiteturas foram desenvolvidas para resolver problemas em diferentes domínios ou casos de uso. Por exemplo, CNN é usado na maior parte do tempo em visão computacional e reconhecimento de imagem, e RNN é comumente usado em problemas relacionados a séries temporais. Os autores informam que as três arquiteturas mais comuns de rede neural profunda são *Convolution Neural Network*, *Autoencoder* e *Restricted Boltzmann Machine*

(RBM).

Pouyanfar et al. (2018) realizou um *survey* sobre algoritmos, técnicas e aplicações de DL. Foi desenvolvida uma adaptação do levantamento dos autores e apresentamos na Tabela 1 a relação das redes de DL e de suas aplicações.

Tabela 1: Redes de Deep Learning e suas aplicações

Redes Deep Learning	Aplicações
<i>Recursive Neural Network</i> (RvNN)	Usa uma estrutura <i>tree-like</i> /Preferido para Processamento de linguagem natural.
<i>Recurrent Neural Network</i> (RNN)	Bom para informações sequenciais / Preferido para Processamento de linguagem natural e processamento de fala.
<i>Convolutional Neural Network</i> (CNN)	Originalmente para reconhecimento de imagem estendido para Processamento de linguagem natural, processamento de fala e visão computacional.
<i>Deep Belief Networks</i> (DBN)	Aprendizagem não supervisionada/ Conexões direcionadas.
<i>Deep Boltzmann Machine</i> (DBM)	Modelo composto de aprendizagem não supervisionada de Máquinas Boltzmann restritas (RBMs) / Conexões não direcionadas.
<i>Generative Adversarial Network</i> (GAN)	Aprendizagem não supervisionada / Framework teórico de jogo.
<i>Variational Autoencoder</i> (VAE)	Modelo gráfico probabilístico de aprendizagem não supervisionada

Fonte: Adaptado de (POUYANFAR et al., 2018)

2.1.1 Redes Neurais Convolucionais

Redes convolucionais profundas trouxeram avanços no processamento de imagens, vídeo, fala e áudio, enquanto as redes recorrentes iluminaram dados sequenciais, como texto e fala (DONG; WANG; ABBAS, 2021) e (KAJAREE DAS, 2017). De acordo com Yunchao e Jiayao (2019), CNN é um tipo de representação de ML e é considerado o melhor algoritmo de ML de reconhecimento de imagem e também o algoritmo de ML mais usado no campo da análise de imagens médicas. LeCun, Bengio e Hinton (2015) informam que o principal motivo é que a CNN corresponde perfeitamente à estrutura bidimensional da imagem e considera essa relação espacial como o valor de entrada. Abaixo estão descritas as camadas que serão utilizadas na arquitetura proposta, seus conceitos e as definições de cada uma delas.

Convolução: Albawi, Mohammed e Al-Zawi (2017) explica que na convolução, em vez de uma conexão completa, são buscadas regiões locais na imagem em vez de na imagem inteira. Em outras palavras, os neurônios ocultos na próxima camada obtêm entradas apenas da parte correspondente da camada anterior. Para Indolia et al. (2018), o vetor de peso, também conhecido como filtro ou kernel, desliza sobre o vetor de entrada para gerar o mapa de características

(*feature maps*). Este método de deslizar o filtro horizontalmente e verticalmente é chamado de operação de convolução.

Pooling (agrupamento): De acordo com Indolia et al. (2018), a localização exata de uma característica torna-se menos significativa uma vez que foi detectada. A ideia principal do Pooling é a redução da amostragem, a fim de reduzir a complexidade para outras camadas. No domínio do processamento de imagens, pode ser considerado semelhante à redução da resolução. O Pooling não afeta o número de filtros. O Maxpooling é um dos tipos mais comuns de métodos de pool. Ele particiona a imagem em retângulos de sub-região e retorna apenas o valor máximo do interior dessa sub-região (ALBAWI; MOHAMMED; AL-ZAWI, 2017).

Não linearidade Albawi, Mohammed e Al-Zawi (2017) A não linearidade pode ser usada para ajustar ou cortar a saída gerada. Esta camada é aplicada a fim de saturar a saída ou limitar a saída gerada. Os autores informam que por muitos anos, sigmóide e tanh foram a não linearidade mais popular e que, recentemente, a Unidade Linear Retificada (ReLU) tem sido usada com mais frequência (INDOLIA et al., 2018).

Conforme Ding, Qian e Zhou (2018), as funções de ativação não lineares trouxeram às redes neurais suas capacidades não lineares, o que significa que as redes neurais podem diferenciar aqueles dados que não podem ser classificados linearmente no espaço de dados atual. Na arquitetura proposta, serão utilizadas as funções ReLU e Sigmóide.

Função ReLU: A função ReLU permite que uma rede obtenha facilmente uma representação esparsa. Mais específico, a saída é zero quando a entrada menor que zero, o que fornece a dispersão na ativação das unidades de neurônios e melhora a eficiência do aprendizado de dados (INDOLIA et al., 2018) e (ALBAWI; MOHAMMED; AL-ZAWI, 2017).

Função Sigmóide: A função sigmóide ajusta o valor de entrada para um intervalo de 0.0 e 1.0. É frequentemente empregada no nível de saída de uma rede neural devido à sua distribuição de valor. Enquanto a função sigmóide raramente é adotada nas redes neurais profundas, exceto o nível de saída desde sua saturação (DING; QIAN; ZHOU, 2018) e (INDOLIA et al., 2018).

Camada Totalmente Conectada, Camada Densa ou Affine: Conforme Indolia et al. (2018), nesta fase, a saída da primeira fase (inclui convolução e agrupamento repetidamente) é alimentada na camada totalmente conectada, e o produto escalar do vetor de peso e do vetor de entrada é calculado para obter a saída final. Cada neurônio recebe os dados de entrada de todos os neurônios da camada anterior (WATAYA et al., 2020). Albawi, Mohammed e Al-Zawi (2017) afirma que a camada totalmente conectada é semelhante à maneira como os neurônios são organizados em uma rede neural tradicional. Portanto, cada nó em uma camada totalmente conectada está diretamente conectado a cada nó na camada anterior e na próxima,

Camada de saída: De acordo com Wataya et al. (2020), nesta camada, os resultados do cálculo da camada oculta anterior são combinados na saída final da rede. Essa camada geralmente é definida como afim ou convolucional; entretanto, o número de neurônios e a função de ativação adequada são determinados de acordo com a tarefa. No caso desta arquitetura, teremos um neurônio de saída que irá definir se a imagem contém ou não câncer.

Função de perda: Essa função quantifica a diferença entre a saída do modelo e a resposta perfeita. Funções de perda bem conhecidas incluem entropia cruzada binária, entropia cruzada categórica e perda quadrática média. Semelhante à camada de saída, a seleção de uma função de perda adequada é limitada dependendo da tarefa particular.

Na Tabela 2, é possível identificar os exemplos camadas de saída e funções de perda e suas finalidades. Para arquitetura proposta, que possui 2 classes, a Entropia Cruzada Binária atuará como função de perda.

Tabela 2: Camadas de saída e funções de perda e suas finalidades

Finalidade	Classificação com 2 classes	Classificação com mais 2 classes
Tipo de camada	Camada Densa	Camada Densa
Número de neurônios	Um	O mesmo que o número de classes
Função de ativação	Função Sigmóide	Função Softmax
Função de perda	Entropia cruzada binária	Entropia cruzada categórica

Fonte: Adaptado de (WATAYA et al., 2020)

2.2 Visão Computacional aplicada à Medicina

De acordo com Dong, Wang e Abbas (2021), visão computacional é um campo interdisciplinar que trata como os computadores podem compreender em alto nível, a partir de imagens ou vídeos digitais. É uma aplicação essencial da inteligência artificial que utiliza computadores e câmeras para substituir o olho humano para o reconhecimento, rastreamento, medição e outros problemas visuais.

Conforme Wiley e Lucas (2018), a visão computacional tem sido estudada por muitas perspectivas. Ele se expande da gravação de dados brutos para técnicas e ideias que combinam processamento de imagem digital, reconhecimento de padrões, ML e computação gráfica. De acordo com os autores, visão computacional é uma combinação de processamento de imagem e reconhecimento de padrões. A saída do processo de Visão por Computador é a compreensão da imagem. O desenvolvimento deste campo é feito adaptando-se a capacidade da visão humana em obter informações.

Para Esteva et al. (2021), a visão computacional tem uma história que abrange décadas de esforços para permitir que os computadores percebam estímulos visuais de forma significativa. A percepção da máquina abrange uma variedade de níveis, desde tarefas de baixo nível, como a identificação de bordas, até tarefas de alto nível, como a compreensão de cenários completos. Os autores informam que o crescimento da visão computacional moderna coincidiu com a geração de grandes quantidades de dados digitais em vários campos científicos e que os avanços médicos recentes têm sido produtivos.

Segundo Fatima e Pasha (2017), em imagens médicas, o Diagnóstico Auxiliado por Computador (CAD, do inglês *Computer Aided Diagnosis*) é uma área dinâmica de pesquisa em rápido

crescimento. Nos últimos anos, foram feitas tentativas significativas para aprimorar os aplicativos de diagnóstico auxiliado por computador, porque os erros nos sistemas de diagnóstico médico podem resultar em tratamentos médicos seriamente enganosos.

Habuza et al. (2021) afirma em seu artigo que a IA para imagens médicas elimina a subjetividade em um procedimento de diagnóstico visual e permite a combinação de imagens médicas com dados clínicos, riscos de estilo de vida e dados demográficos. Mas pondera que o projeto de um classificador automático ideal deve incorporar tanto o conhecimento especializado sobre uma doença quanto as técnicas de visão computacional de última geração.

Conforme Aeffner et al. (2019), o termo “análise de imagens” foi reservado para a disciplina específica que visa obter informações significativas de imagens de forma objetiva e reproduzível. A origem da análise de imagens com ferramentas objetivas é quase tão antiga quanto a própria microscopia. Os autores informam que embora tenha havido muitos avanços incrementais ao longo dos séculos, a análise de imagens permaneceu praticamente inalterada até o advento da imagem digital e da análise computadorizada na segunda metade do século passado.

De acordo com Chang et al. (2016), as imagens podem ser obtidas de várias maneiras: sistemas de imagens médicas, scanners 3D a laser, microscópios ou satélites. Há um aumento acentuado nos aplicativos de informática de imagem, visto que houveram avanços simultâneos em plataformas de imagem, disponibilidade de dados devido às mídias sociais e análise de big data. Os autores informam que scanners, como imagens de ressonância magnética ou tomografia computadorizada, frequentemente usados em radiologia diagnóstica, produzem dados de imagem digital e que, mais recentemente, a patologia digital está contando com imagens de slides inteiros, sistemas automatizados de microscopia que registram imagens de slides, em múltiplas resoluções.

Arora, Tyagi e Manzar (2020) complementam informando que na área médica, várias tecnologias surgiram nos sistemas de diagnóstico em que mais comumente disponíveis são os raios X, a tomografia computadorizada e a ressonância magnética, das quais a ressonância magnética é a mais útil e inofensiva por não haver radiação e que isso se baseia principalmente no campo magnético e nas ondas de rádio. As imagens fornecem informações detalhadas sobre as áreas onde são encontradas anormalidades, como o tipo de tumor, a posição e o tamanho.

De acordo com Aeffner et al. (2019), com o advento da imagem de lâmina inteira na patologia digital houve o avanço do exame de tecido auxiliado por computador por meio da análise de imagem digital. Lâminas digitalizadas podem ser facilmente anotadas e analisadas por meio de uma variedade de algoritmos. Para Srinidhi, Ciga e Martel (2021), imagens histopatológicas¹ contêm informações fenotípicas ricas que podem ser usadas para monitorar os mecanismos subjacentes que contribuem para a progressão da doença e os resultados de sobrevida do paciente.

Sobre as análises das imagens, Madabhushi e Lee (2016) informa que as medições baseadas em área incluem as avaliações mais básicas como quantificar as áreas de uma determinada mancha, por exemplo, mancha química ou imunohistoquímica (IHC), a área de vacúolos de

¹Estudo de como uma doença específica afeta um conjunto de células.

gordura ou outros eventos presentes em uma lâmina. Conforme Srinidhi, Ciga e Martel (2021), o exame e a interpretação de seções de tecido coradas com hematoxilina e eosina (H&E) por patologistas anatômicos é um componente essencial na avaliação da doença. Além de fornecer informações diagnósticas, as informações fenotípicas contidas nas lâminas histológicas podem ser utilizadas para o prognóstico.

2.3 Detecção de Câncer por Imagem

Conforme a definição do Instituto Nacional de Câncer - INCA (2021), câncer é definido como o crescimento desordenado de células, que podem invadir tecidos adjacentes ou órgãos a distância. De acordo com Albino de Queiroz et al. (2021), o câncer é a segunda principal causa de morte no mundo, sendo responsável por uma estimativa de 9,6 milhões de mortes, correspondendo a um sexto das mortes em 2018 e que, no Brasil estima ter 626.030 novos casos de câncer em 2020. De acordo com o INCA (2021), existem mais de cem diferentes tipos de doenças malignas, como Anal, Bexiga, Boca, Colo do útero, Corpo do útero, Esôfago, Estômago, Fígado, Infantojuvenil, Intestino, Laringe, Leucemia, Linfoma de Hodgkin, Linfoma não Hodgkin, Tireoide, Ovário, Pâncreas, Pele melanoma, Pele não melanoma, Pênis, Próstata, Pulmão, Sistema Nervoso Central, Testículo e Mama.

Sobre o câncer de mama, alvo deste estudo, é um tumor maligno que ocorre no epitélio glandular da mama, muitas vezes, pelo processo de crescimento celular errado. Quando isso ocorre, o acúmulo de células geralmente forma uma massa de tecido chamada protuberância, nódulo ou tumor (Li et al., 2018). Os autores ainda informam que existem várias técnicas de obtenção de imagem que podem ser utilizadas conforme Figura 1.



Fonte: Adaptado de Li et al. (2018)

Segundo Kim et al. (2017) existem muitos tipos histopatológicos de câncer de mama, sendo o tipo mais frequente o carcinoma invasivo, desta forma a expressão imuno-histoquímica do receptor de estrogênio (ER), receptor de progesterona (PR) e o receptor de fator de crescimento epitelial humano 2 (HER2) tem sido amplamente utilizada para prever o prognóstico do câncer de mama e para fornecer estratégias terapêuticas. Esses marcadores podem ser identificados

através de técnicas convencionais de imuno-histoquímicas (IHC) (OLIVER; JAMUR, 2010).

De acordo com Salvio e Marques (2006), as imagens são difíceis de se distinguir nas colorações de rotina, em lâminas coradas pela hematoxilina eosina (H&E). Neste caso pode-se utilizar técnicas de imuno-histoquímica como utilização da 3-3' diaminobenzidrina (DAB) como cromógeno. O DAB atua como cromógeno, que na presença da enzima peroxidase produz um precipitado de cor marrom, revelando desta forma os tecidos imunomarcados (OLIVEIRA; OLIVEIRA; PINHEIRO, 2011).

2.3.1 Patologia Digital e uso de Deep Learning

Para realizar as análises imunohistoquímicas dessas imagens é geralmente utilizado o escore histológico (HSCORE), na qual é observada a porcentagem da área examinada e a intensidade percebida da coloração. Deste modo a habilidade e experiência do observador podem influenciar os resultados obtidos. Em seu trabalho Fuhrich, Lessey e Savaris (2013), utilizando o software ImageJ, em um processo que o autor denomina Digital HSCORE (D-HSCORE), compara análise realizada por um novo pesquisador sem experiência na avaliação HSCORE com um especialista na área com mais de 20 anos de experiência em leitura de análises imunohistoquímicas e conclui que o D-HSCORE realizado por um pesquisador inexperiente tem alta correlação com o HSCORE tradicional realizado por um especialista.

O auxílio de softwares tem servido de apoio para a patologia digital, e os avanços da Inteligência Artificial tem apresentado grande potencial nesta área. De acordo com Salvi et al. (2021), frameworks de DL tornaram-se rapidamente a principal metodologia de análise de imagens médicas e que os algoritmos de DL são ideais para a área de patologia digital devido a sua capacidade de aprendizado e às vantagens em lidar com padrões complexos. Ainda sobre patologia digital, os autores informam que as CNNs tem sido amplamente utilizadas, onde lâminas histológicas são digitalizadas para produzir imagens de alta resolução, graças a scanners digitais de lâmina inteira. Segundo Echle et al. (2021), os avanços de DL permitiram a extração de informações anteriormente ocultas diretamente de imagens histológicas de rotina de câncer, fornecendo informações potencialmente úteis clinicamente

2.3.2 Biglycan

De acordo com (QUEZADA-MARÍN et al., 2020), um biomarcador é uma característica que é objetivamente medida e avaliada como um indicador de processos normais/patogênicos ou respostas terapêuticas farmacológicas. É qualquer substância, estrutura ou processo que pode ser medido no corpo e influenciar ou prever a incidência de um resultado ou doença.

O Biglycan (BGN) é um membro da família de pequenos proteoglicanos ricos em leucina, consiste em uma proteína principal de 38 kDa e duas cadeias de condroitina / sulfato de dermatan (YANG et al., 2015). Atualmente, o biglycan tem muitas funções propostas. Cada função

proposta depende do tecido ou órgão em questão (WADHWA et al., 2004). Estudos clínicos demonstraram que a regulação positiva da BGN está associada a um mau prognóstico em pacientes com vários tipos de câncer sólido (ZHAO et al., 2020).

Em um estudo feito por Zhao et al. (2020) concluíram que a alta expressão de RNA mensageiro (mRNA) de BGN foi indicada estar associado a um desfecho clínico ruim no prognóstico de pacientes com câncer de bexiga, carcinoma de células escamosas do pulmão e cistoadenocarcinoma seroso do ovário. No mesmo estudos, os autores demonstraram haver alta expressão de mRNA de BGN em pacientes com câncer de mama comparados com pacientes saudáveis. Outros estudos vem demonstrando como o biglycan pode estar influenciando o comportamento das células tumorais e por isso resolveu-se estudar melhor a relação desta proteína biglycan e o câncer de mama como um potencial biomarcador para esta doença.

Os biomarcadores prognósticos permitem a categorização dos pacientes de acordo com seu risco de progressão da doença ou morte e, portanto, podem ser usados para ajustar a intensidade do tratamento para pacientes individuais. Por outro lado, os biomarcadores preditivos permitem que um determinado tratamento direcionado seja escolhido para um grupo específico de pacientes (ECHLE et al., 2021). Os biomarcadores diagnósticos destinam-se a ajudar os patologistas a estabelecer um diagnóstico específico; marcadores preditivos indicam a probabilidade de os pacientes se beneficiarem de uma terapia específica; e marcadores prognósticos determinam os resultados dos pacientes, na ausência de tratamentos específicos (QUEZADA-MARÍN et al., 2020).

3 TRABALHOS RELACIONADOS

A computação como ferramenta de apoio para medicina é amplamente utilizada há vários anos. Desde criação de sensores, robôs para exames e cirurgias bem como para o auxílio de diagnósticos de doenças.

De acordo com Lu et al. (2018), um dos desafios enfrentados é a detecção de grandes quantidades de imagens, realizadas, muitas vezes, manualmente, o que acarreta em um grande espaço de tempo demandado e também não há possibilidades de garantia de precisão, nem mesmo para especialistas.

A etiologia² do câncer de mama ainda não é totalmente compreendida. A maneira eficiente de melhorar a precisão da detecção precoce é combinar diferentes métodos de imagem, como raio-x (mamografia), ultrassom e imagem por ressonância magnética (MRI). A análise histopatológica desempenha um papel importante na detecção de metástases no linfonodo sentinela, no contexto da extensão da disseminação do câncer para o estadiamento do câncer de mama (Lu et al., 2018).

Realizou-se duas etapas para conduzir o processo da busca dos trabalhos relacionados. O primeiro é a criação de uma string de pesquisa, enquanto o segundo envolve a seleção das fontes. No primeiro, o processo envolve a definição de palavras-chave de pesquisa e a definição do escopo da pesquisa (GUBERT; COSTA; RIGHI, 2019). As palavras-chave de pesquisa foram definidas após a leitura de artigos na área de interesse e a separação de termos, sinônimos e acrônimos que melhor definiram o objeto de pesquisa. Seguindo o que foi definido por Kitchenham (2004), esses termos foram combinados usando operadores booleanos para melhorar o desempenho da string de pesquisa.

("Machine Learning"OR "Deep Learning"OR "Convolutional neural network ") AND (Cancer OR "Brest Cancer"OR "Cancer biomarker")

Para a seleção das fontes, foram obtidos estudos em bases de dados eletrônicos, através de pesquisas usando as palavras-chave da string de pesquisa. Os bases utilizadas para a pesquisa foram: ACM Digital Library³; IEEE Xplore Library⁴; ScienceDirect⁵; Springer Link⁶; e Google scholar⁷. Essas bases de dados foram escolhidas porque constituem uma amostra significativa de bases de dados de referência para área de pesquisa e fornecem artigos completos de periódicos e anais de eventos das mais importantes conferências que envolvem os temas de pesquisa relacionados, podendo ser observado na Tabela 3.

Segundo Srinidhi, Ciga e Martel (2021), as imagens histopatológicas contêm informações

²Ramo da Medicina que estuda as causas das doenças.

³<https://dl.acm.org/>

⁴<https://ieeexplore.ieee.org/>

⁵<https://www.sciencedirect.com/>

⁶<https://link.springer.com/>

⁷<https://scholar.google.com.br/>

que podem ser usadas para monitorar os mecanismos subjacentes que contribuem para a progressão da doença e os resultados de sobrevida do paciente. Recentemente, Deep Learning se tornou a principal escolha metodológica para analisar e interpretar imagens histológicas.

Ainda, de acordo com Srinidhi, Ciga e Martel (2021), a introdução de scanners de lâmina inteira na década de 1990 criou um interesse na aplicação de análise de imagem e técnicas de ML à histopatologia, além de algoritmos criados originalmente para diagnóstico em radiologia foram adaptados com sucesso para uso em patologia digital. Esses avanços levaram a um crescente interesse no uso de biomarcadores extraídos de imagens de patologia digital para medicina de precisão, particularmente em oncologia. Mais tarde, Krizhevsky, Sutskever e Hinton (2012) mostraram que as CNNs poderiam superar as abordagens de ML anteriores, classificando 1,2 milhão de imagens de alta resolução no concurso ImageNet ILSVRC-2010 em 1000 classes diferentes.

(Shabbeer Basha et al., 2018) propõe o RCCNet, uma arquitetura baseada em CNN para a classificação de câncer de cólon utilizando a técnica de histologia. O principal objetivo dessa rede é manter o modelo da CNN o mais simples possível e demonstrou ser compacto e otimizado para áreas histológicas de baixa resolução, utilizando 7 camadas simples com 1.512.868 parâmetros treináveis. Os experimentos utilizaram o dataset "CRCHistoPhenotypes", que é um conjunto de dados histológicos de câncer de cólon e os resultados do modelo proposto são comparados com cinco modelos CNN de última geração em termos de precisão, pontuação média ponderada de F1 e tempo de treinamento. O método proposto alcançou uma acurácia de classificação de 80,61% e média ponderada de pontuação F1 de 0,7887.

Outro exemplo de uso de rede de DL foi o de Li et al. (2018), realizado para identificação do câncer gástrico através do framework GastricNet. Em sua proposta, os autores recortaram áreas com tamanho 224 x 224 de imagens gástricas para formar o conjunto de treinamento, uma vez que a resolução originais são de 2048 x 2048, muito grande para a rede DL processar diretamente. GastricNet adota arquiteturas diferentes para camadas rasas e profundas, ou seja, módulo multi-escala para camadas superficiais e módulo Rede em Rede para camadas profundas. Em seus resultados experimentais, o GastricNet superou vários frameworks de DL conhecidos não apenas na classificação patch-based, mas também atingiu uma precisão de classificação média de 100% para a classificação sliced-based.

De acordo com Chang et al. (2017), CNN tem sido amplamente utilizada para construir modelos de classificação de imagens que permitem aos pesquisadores transferir o modelo de aprendizagem pré-treinado para outras classificações. Os autores propõem um método de aprendizagem de transferência para detectar câncer de mama usando imagens histopatológicas com base no modelo Inception v3 do Google, que foram inicialmente treinados para a classificação de imagens não médicas. O modelo treinado realizou uma classificação com acurácia de 0,83 para a classe benigna e 0,89 para a classe maligna.

Li et al. (2018), utilizaram CNN para classificar anormalidades, benignas ou malignas, em imagens mamográficas com base no banco de dados mini-MIAS. Para melhorar o desempenho,

utilizaram métodos de pré-processamento de imagem contendo recorte, normalização de contraste global, aumento, equalização de histograma local e pré-processamento de equilíbrio. O modelo proposto pelos autores, chamado de CNN4-d, obteve melhor desempenho com uma precisão de treinamento de 93,43% e precisão de teste de 89,05%. O experimento mostrou também que a estratégia de dropout em camadas totalmente conectadas diminui o nível de *overfitting* e obtiveram melhor desempenho na classificação.

Yeh e Chan (2018) trabalharam com imagens por meio do método de Tomossíntese digital de mama (DBT). Contudo os autores informaram que este método gera inúmeras imagens, criando uma grande carga de trabalho para os radiologistas e portanto, era necessário desenvolver um sistema CAD automático para análise de imagens DBT. Em seu estudo, os autores compararam CAD baseado em recursos e CAD baseado em CNN para classificação de câncer de mama. As taxas de precisão (desvio padrão) do CAD baseado em recursos e em CNN para a classificação do câncer de mama foram 74,85% (0,122) e 87,12% (0,035), respectivamente. O valor T foi -6,229 e o valor P foi $0,00 < 0,05$, o que indicou que o CAD baseado em CNN superou significativamente o CAD baseado em recursos.

Uma outra abordagem foi utilizada por Turkki et al. (2019), que empregou CNN para prever o resultado do paciente com base apenas na aparência visual do tecido do câncer de mama, sem quaisquer suposições anteriores como quantificação intermediária de estruturas de tecido, como tipos e estados específicos de células (por exemplo, células mitóticas, células pleomórficas, células imunes) ou estruturas e entidades de tecido (dutos, necrose, vasos). Neste estudo é utilizado o Tissue Microarray, que é uma composição de várias biópsias de mama agrupadas em um bloco de parafina, utilizando uma análise com hematoxilina e eosina.

Em seu trabalho, Mahmood et al. (2020) apresentou um método de detecção de células mitóticas em vários estágios baseado na CNN da região mais rápida (Faster R-CNN) e CNNs profundos. Foram utilizados para os experimentos, dois conjuntos de dados abertos da Conferência Internacional sobre Reconhecimento de Padrões (ICPR 2012 e ICPR 2014 (MITOS-ATYPIA-14) de histopatologia de câncer de mama. Os resultados experimentais mostraram que o método atingiu os resultados mais avançados de 0,876 precisão, 0,841 recall e 0,858 medida F1 para o *dataset* ICPR 2012 e 0,848 precisão, 0,583 recall e 0,691 medida F1 para o *dataset* ICPR 2014.

Outro trabalho que utilizou CNN em contexto de câncer da mama foi em Zeiser et al. (2021), que propôs um modelo para fornecer uma segmentação refinada e multiclasse de Whole Slide Imaging (WSI) para câncer de mama. A metodologia é dividida em quatro módulos: pré-processamento, detecção de Região de Interesse (do inglês *Region of interest* - ROI), Amostragem de ROI e segmentação de região. Os datasets utilizados pelo trabalho foram: The Grand Challenge on Breast Cancer Histology Images (BACH), The Cancer Genome Atlas Program (TCGA) e HASHI. O modelo combinou duas CNNs em cascata para identificar ROIs e para analisar as regiões selecionadas. De acordo com os autores, os resultados indicaram a possibilidade de utilização do modelo como um segundo sistema de triagem.

Tabela 3: Trabalhos Relacionados

Trabalho	Técnica de ML	Tipo de Câncer	Detecta Câncer	Origem de imagem
(Shabbeer Basha et al., 2018)	CNN	Cólon	NÃO, faz classificação	Dataset CRCHistoPhenotypes - núcleos histológicos
Li et al. (2018)	CNN	Gástrico	SIM	Dataset de lâminas gástrica BOT
Yeh e Chan (2018)	CNN	Mama	SIM	Tomossíntese digital da mama (Digital breast tomosynthesis - DBT)
Li et al. (2018)	CNN (CNN-4d)	Mama	SIM	Mamograma (Raio - X)
Turkki et al. (2019)	CNN (VGG-16)	Mama	SIM	Tecido Microarray (TMA) de tecido mamário
Mahmood et al. (2020)	Faster-CNN	Mama	SIM	Datasets ICPR 2012 e ICPR 2014 (MITOS-ATYPIA-14)
Zeiser et al. (2021)	CNN	Mama	SIM	BACH, TCGA e HASHI.

Fonte: Elaborado pelo autor

4 ARQUITETURA DE *DEEP LEARNING* PARA DETECÇÃO DE CÂNCER DE MAMA COM BGN

O modelo proposto neste trabalho, visa, a partir de um conjunto de imagens histológicas, disponibilizadas pelo serviço de patologia do Hospital das Clínicas de Porto Alegre, criar um conjunto de dados (*dataset*) totalmente novo a partir dessas imagens, utilizando o Digital HS-CORE, uma forma menos subjetiva para identificar a intensidade das cores, e gerar uma CNN para treinar um modelo que verifique se existe diferença entre a expressão do Biglycan entre tecidos com e sem câncer de mama. Além do uso do próprio *dataset* criado, a arquitetura proposta se difere por utilizar imagens histológicas, com o biomarcador BGN, através da intensidade da coloração do DAB, utilizando a deconvolução das cores, com o objetivo de identificar se a expressão de um potencial biomarcador é diferente entre tecidos mamários com e sem malignidade, contribuindo para a literatura.

Considerando os aspectos citados a priori, o modelo proposto neste trabalho visa verificar se a expressão da proteína Biglycan é diferente entre tecidos mamários com e sem malignidade. A associação do DL e da expressão da proteína BGN, através da intensidade da coloração do DAB, utilizando a deconvolução das cores é nova; há a combinação entre análise de imagem semi-supervisionada e quantificação da expressão de uma proteína que ainda não foi validada como biomarcador para o câncer de mama. Histologicamente, há uma diferença bem estabelecida entre o tecido com e sem câncer de mama, a qual é realizada nos laboratórios de patologia. Apresentar estas imagens para análise por CNN não irá determinar a expressão da proteína, uma vez que há um desarranjo da arquitetura das células pelo câncer.

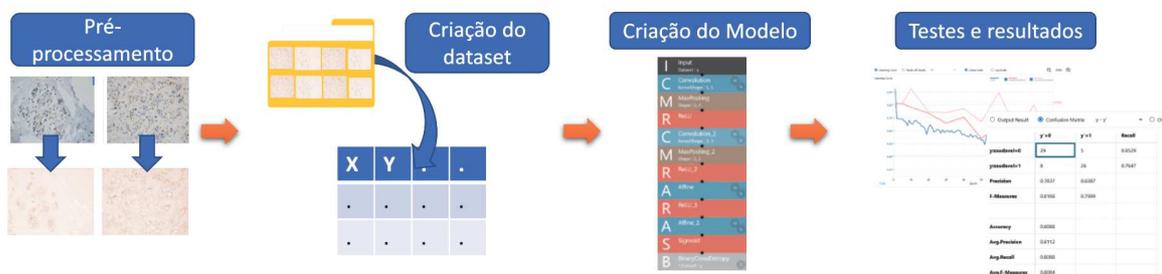
4.1 Arquitetura Proposta

A arquitetura proposta utilizou um conjunto de dados criado a partir de imagens de pacientes do Hospital de Clínicas de Porto Alegre. O *dataset* foi utilizado para testar o modelo da CNN que irá classificar a amostra do paciente - com ou sem câncer de mama. O modelo de CNN foi utilizado para implementar DL tendo em vista que a sua aplicação foi a mais recomendada dentro da literatura.

Para facilitar o entendimento, a Figura 2 apresenta uma visão geral do fluxo de trabalho das ações executadas desde o pré-processamento das imagens até a obtenção dos resultados. Na arquitetura proposta, o uso de DL, através de uma proposta de uma arquitetura de CNN, possibilita o processamento de dados do *dataset* criado com informações de tecidos saudáveis e com câncer. Inicialmente, esses dados serão usados para classificar se a imagem contém câncer ou não. Em seguida, será avaliada também a intensidade para do biomarcador BGN com a finalidade de alimentar os modelos de ML. Inicialmente serão realizados os testes sem alterar ou recortar as imagens recebidas, de forma a verificar o desempenho dessa CNN. As classificações feitas ficarão disponíveis através de uma aplicação para que seja possível fazer

consultas e monitorar as informações médicas de um determinado paciente.

Figura 2: Fluxo das atividades executadas



Fonte: Elaborado pelo autor

Inicialmente, as imagens recebidas possuem uma resolução de 4800 x 3600 pixels. As mesmas não sofrerão recortes para os treinamentos do modelo, contudo serão redimensionadas para que suas proporções estejam de acordo com a criação do dataset e o treino do mesmo. Após os primeiros testes, será verificada a necessidade de aumentar ou diminuir a resolução das mesmas e também se será necessário utilizar alguma técnica para aumento da quantidade das imagens.

De acordo com Albawi, Mohammed e Al-Zawi (2017), CNN é uma das redes neurais profundas mais populares e que seu nome deriva da operação matemática linear entre matrizes chamada convolução. A CNN tem várias camadas; incluindo camada convolucional, camada de não linearidade, camada de pooling e camada totalmente conectada. Os autores ainda informam que a CNN tem excelente desempenho em problemas de ML, principalmente os aplicativos que lidam com dados de imagem.

4.2 Dataset e Pré-processamento

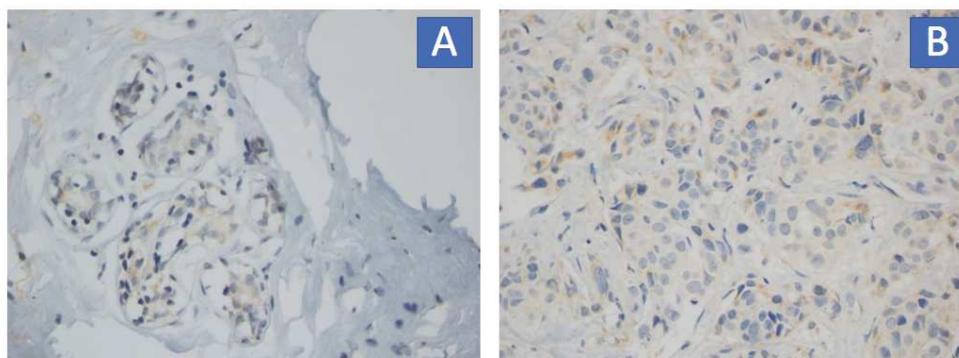
O conjunto de dados foi construído a partir de imagens disponibilizadas pelo serviço de patologia do Hospital de Clínicas de Porto Alegre (HCPA), composto por imagens oriundas de cortes histológicos com e sem tumor de câncer de mama. Este estudo foi inicialmente aprovado pelo Comitê de Ética em Pesquisa do Hospital de Clínicas de Porto Alegre (2019/0337) CAAE 15329119.9.0000.5327 para ser analisado utilizando a análise digital das imagens pelo método de D-HSCORE, sendo este estudo uma colaboração entre as duas instituições.

O cenário do estudo de caso-controle onde as imagens foram obtidas apresenta 36 tecidos fixados em formalina e embebidos em parafina que foram captadas para análise de arquivos patológicos. Os tecidos mamários benignos (n = 9; controle) e câncer (n = 27) foram analisados por imuno-histoquímica BGN monoclonal antibody (M01), clone 4E1-1G7(Ab60037), na diluição 1:100 a pH 9. A análise imuno-histoquímica foi realizada de forma independente por dois analistas em microscopia de luz e 10% dos resultados finais foram revisados por um pa-

tologista certificado para confirmar o diagnóstico. De acordo com os patologistas, as amostras onde as células apresentam formatos anárquicos (não redondas), são células que contêm câncer.

Do total, obteve-se 336 imagens rotuladas, compondo um conjunto de treinamento de 268 imagens, um conjunto de teste inicial de 68 imagens. Os profissionais do HCPA rotularam as imagens como saudáveis (controle) e câncer (caso). A Figura 3 ilustra as imagens de cada classe mencionada no conjunto de dados.

Figura 3: Exemplo das imagens disponibilizadas pelo HCPA, sendo A - Saudável e B - Câncer



Fonte: Elaborado pelo autor

A expressão da proteína BGN foi obtida por técnica de imuno-histoquímica utilizando o anticorpo BGN monoclonal antibody (M01) (clone 4E1-1G7 - Abnova Corporation, camundongo anti-human). As imagens foram obtidas com uma objetiva UPLanFI de 100x (resolução: 2,75 mm), com um tamanho de 4800 x 3600 pixels, sob condições padrão.

Do total, obteve-se 336 imagens rotuladas, sendo 203 com câncer e 133 saudáveis. O número de imagens no conjunto de treinamento e conjunto de teste das imagens é mostrado na Tabela 4. A proporção utilizada para treino e teste foi a mesma de (HUANG; LIN, 2020), (KER et al., 2019).

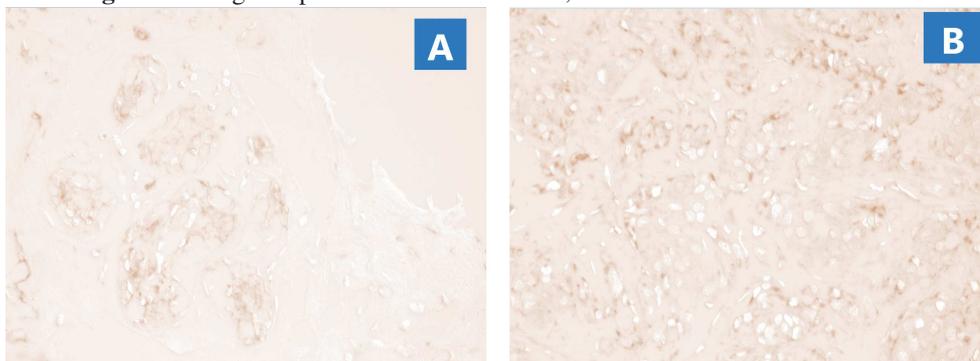
Tabela 4: Dataset - Conjunto de imagens de treino e teste

Rótulo	Total	Treinamento	Teste
Câncer	203	171	32
Saudável	133	97	36

Fonte: Elaborado pela autor.

O conjunto de dados neste trabalho foi construído para ser usado em rede neural convolucional para a classificação de amostras saudáveis e de câncer. Assim como Huang e Lin (2020), as imagens originais foram redimensionadas para unificar o seu formato, utilizou-se a proporção de 128 x 128. O método de processamento da imagem escolhido foi o Resize, que apresentou melhores resultados nos testes preliminares. Também foram testados os modelos Trimming e Padding. Além da obtenção dos melhores resultados, o método Resize foi escolhido pois não realizava nenhum recorte ou deformação na imagem, o que poderia acarretar em perda de dados.

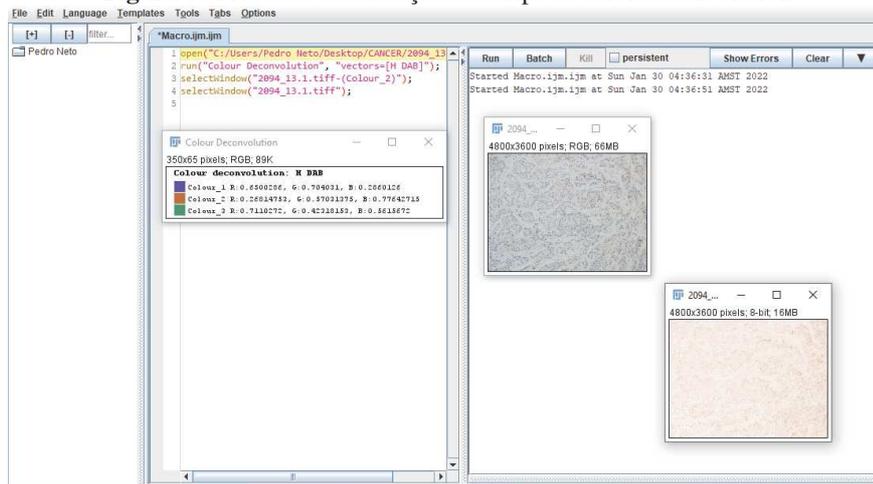
Figura 4: Imagens após Color Deconvolution, sendo A - Saudável e B - Câncer



Fonte: Elaborado pelo autor

Após os primeiros testes, foi solicitado pela equipe de patologistas que fosse realizado o processo de *color deconvolution* para que as imagens estivessem em tons de marrom, para que o *dataset* refletisse o resultado a partir da proteína BGN, pois diminuiria as chances da CNN estar avaliando somente as diferenças na arquitetura de células e tecidos entre o tecido saudável e com câncer. Uma situação semelhante também foi relatada por Ker et al. (2019). O *color deconvolution* foi realizado pelo software Image J, nome da versão anterior do programa Fiji, e o mesmo software utilizado pelo serviço de patologia que forneceu as imagens e podem ser verificados na Figura 4.

Figura 5: Processo de execução de script do color deconvolution



Fonte: Elaborado pelo autor

Durante esta etapa, inicialmente foi realizado o processo de deconvolução de forma manual, alterando as imagens uma por uma, o que demonstrou demandar muito tempo para finalizar toda a tarefa. Para aumentar o desempenho desta ação, foi criado um script, através de captura de macros, para que o software Fiji realizasse, de forma automatizada, a coloração usando

Figura 8: Implementação do Modelo Proposto

Fonte: Elaborado pelo autor

um processador Intel Core i7 3.4 GHz, 8 GB DDR RAM, sem GPU, tendo sua execução no processador local do equipamento.

Segundo Erickson et al. (2017), o aprendizado de máquina supervisionado tem esse nome porque são necessários exemplos de cada classe a ser aprendida. Uma maneira popular de estimar a precisão de um sistema de ML quando há um conjunto de dados limitado é usar a técnica de validação cruzada, onde, primeiro seleciona-se um subconjunto de exemplos para treinamento e designa-se os exemplos restantes a serem usados para teste. Para avaliar a arquitetura proposta, realizou-se, inicialmente, o treinamento do modelo utilizando uma divisão de dados para treino e teste em uma proporção de 80% para treino e 20% para teste utilizando o `validation_split`. Logo após, foi utilizada validação cruzada, através da técnica K-Fold, que se refere ao número de grupos em que uma determinada amostra de dados deve ser dividida. Quando um valor específico para k é escolhido, ele pode ser usado no lugar de k na referência ao modelo. Neste caso, como $k=5$ tornando-se validação cruzada de 5 vezes. Os blocos foram ajustados para 150 épocas com um tamanho de lote de 68.

4.4 Métricas de Avaliação

Conforme Ker et al. (2019), para avaliar a precisão do desempenho da rede, utiliza-se métricas comuns na avaliação de ML: Recall, Precisão, Pontuação F1 e Acurácia, conforme abaixo:

$$R(\text{Recall}) = \frac{TP}{TP + FN}$$

$$P(\text{Precisao}) = \frac{TP}{TP + FP}$$

$$F1 = \frac{2 * P * R}{P + R}$$

$$ACC(\text{Acuracia}) = \frac{TP + TN}{TP + TN + FP + FN}$$

Fonte: (KER et al., 2019)

Conforme as fórmulas apresentadas, TP (verdadeiro positivo, do inglês, *True Positive*) é o número de casos positivos que são classificados como positivos. Respectivamente, TN (*True Negative*), FN (*False Negative*) e FP (*False Positive*) representam os números de verdadeiros negativos, falsos negativos e falsos positivos, respectivamente. ACC refere-se à Acurácia. O Recall representa a porcentagem de amostras positivas que estão classificadas corretamente, o que é mais clinicamente relevante, conforme Li, Wu e Wu (2019) e F1 avalia a precisão obtida através do conjunto de validação. A matriz de confusão é uma tabela que permite a visualização do desempenho do algoritmo

Também foi utilizado para avaliar o desempenho do modelo a curva ROC Receiver Operating Characteristic (ROC), uma representação gráfica da performance de um modelo de dados quantitativos segundo sua taxa de sensibilidade (fração dos verdadeiros positivos) e a fração dos falsos positivos (1-especificidade) (LI; WU; WU, 2019) (POLO et al., 2020). Com a curva ROC, podemos obter a Área Sob a Curva ROC (do inglês, Area Under the Curve - AUC). A AUC é usada para medir o quão bem um modelo pode diferenciar classes, ou seja, quanto maior o valor da AUC, melhor o desempenho do modelo (ZEISER et al., 2021). A curva ROC mostra o trade off entre a fração positiva verdadeira (TPF) e a fração falsa positiva (FPF) conforme uma mudança no critério de positividade (HAJIAN-TILAKI, 2013).

$$\text{TruePositiveFraction} = \frac{TP}{TP + FN}$$

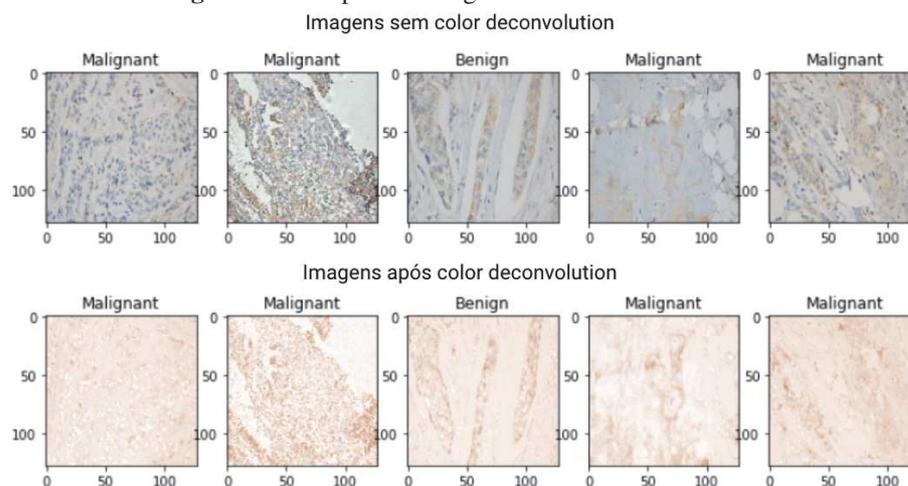
$$\text{FalsePositiveFraction} = \frac{FP}{FP + TN}$$

Fonte: (HAJIAN-TILAKI, 2013)

5 RESULTADOS

Inicialmente foi realizado o treinamento e teste com as imagens sem aplicar o deconvolução de cores no modelo BGNDL, para que seja feita a comparação dos resultados com as imagens que receberam esse tratamento. Desta forma, verificar se existe se o modelo diferencia os dois tipos diferentes de dataset e quais são os resultados obtidos após os treinos desses conjuntos. Exemplos de imagens dos dois conjuntos podem ser observados na Figura 9.

Figura 9: Exemplos de imagens dos datasets utilizados

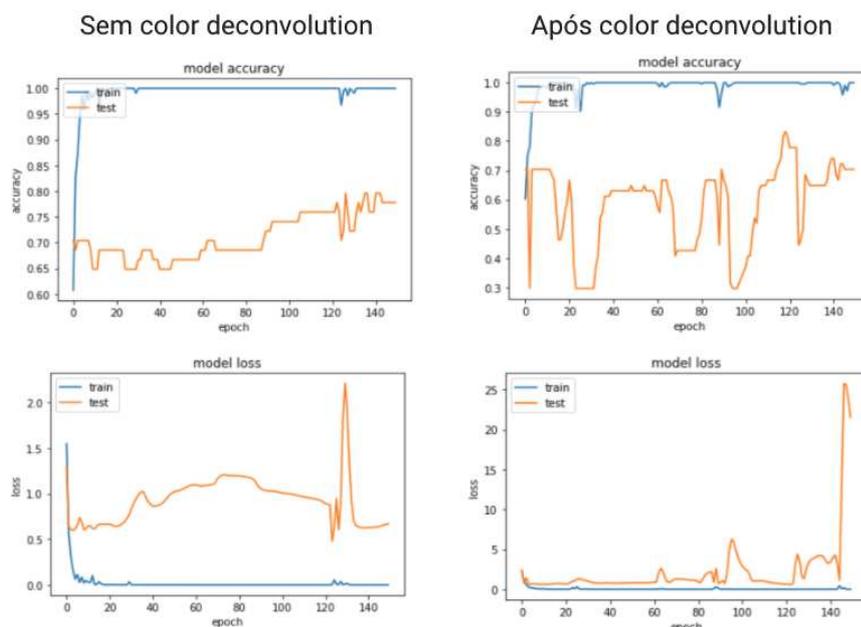


Fonte: Elaborado pelo autor

Em um primeiro momento, foi realizado o treinamento e teste utilizando a função `validation_split` da biblioteca Scikit-learn. É possível observar na Figura 10 os gráficos de acurácia e perda do modelo. Do lado esquerdo, os gráficos do modelo com o dataset das imagens sem receber a deconvolução de cores e do lado direito com o dataset com as imagens após deconvolução de cores. De acordo com Ponti e Costa (2017), mostra o quão longe o estimador está da predição ideal e portanto quantifica o “custo” ou “perda” ao aceitarmos a predição gerada pelos parâmetros atuais do modelo.

Observou-se muito ruído nos gráficos, além de uma lacuna entre as linhas de treinamento e a validação. Segundo Thanapol et al. (2020), o treinamento da CNN geralmente requer um grande número de amostras para um resultado satisfatório e que a obtenção de um número suficiente de amostras pode não ser possível. Para isso, foi realizado a aplicação da validação cruzada, utilizando a técnica K-Fold para realizar um novo treinamento. Com a validação do KFold é possível verificar se o modelo funciona para diferentes condições aleatórias, deste modo, podemos verificar se o modelo funciona bem o tempo todo, ou apenas algumas vezes. Após experimentos, chegou-se ao valor 5 para o tamanho do K-Fold, conforme Figura 11.

Após a realização da validação cruzada com a técnica K-Fold, realizou-se o processo de verificação de predição, utilizando o modelo treinado. Foram utilizadas 68 imagens nunca vistas

Figura 10: Acurácia e Perda do modelo - Imagens sem deconvolução de cores

Fonte: Elaborado pelo autor

Figura 11: Exemplo de Validação-Cruzada K-Fold = 5

Teste	Treino	Treino	Treino	Treino
Treino	Teste	Treino	Treino	Treino
Treino	Treino	Teste	Treino	Treino
Treino	Treino	Treino	Teste	Treino
Treino	Treino	Treino	Treino	Teste

Fonte: Elaborado pelo autor

antes pelo modelo, 32 rotuladas como câncer e 36 rotuladas como saudável. Os resultados obtidos podem ser observados na Figura 12. A matriz de confusão é uma tabela informativa utilizada para analisar os erros e confusões entre diferentes cenas e é obtida contando cada classe de classificação correta e incorreta (CHENG et al., 2018). A Acurácia que, segundo, Wataya et al. (2020), mostra o desempenho do modelo de uma maneira geral, apresentou o valor de 72% para o dataset sem o color deconvolution e 93% para o dataset com o color deconvolution. A precisão, que de acordo com Cheng et al. (2018) é definida como o número de imagens classificadas corretamente, independentemente da classe a que pertencem, dividido pelo número total de imagens, apresentou uma média de 77% e 93% para os dataset sem e com o color deconvolution, respectivamente.

A média de *Recall* - o número de vezes que uma classe foi predita corretamente (TP) dividido pelo número de vezes que a classe aparece no dado de teste (FN), teve média de 73%

Figura 12: Matriz confusão e métricas - modelo BGNDL

Dataset sem Color Deconvolution							
T E S T E	REAL			Acurácia	Precisão	Recall	F1-Score
		Benigno	Maligno	72%	77%	73%	72%
	Benigno	19	17				
Maligno	2	30					
Dataset após Color Deconvolution							
T E S T E	REAL			Acurácia	Precisão	Recall	F1-Score
		Benigno	Maligno	93%	93%	93%	93%
	Benigno	34	2				
Maligno	3	29					

Fonte: Elaborado pelo autor

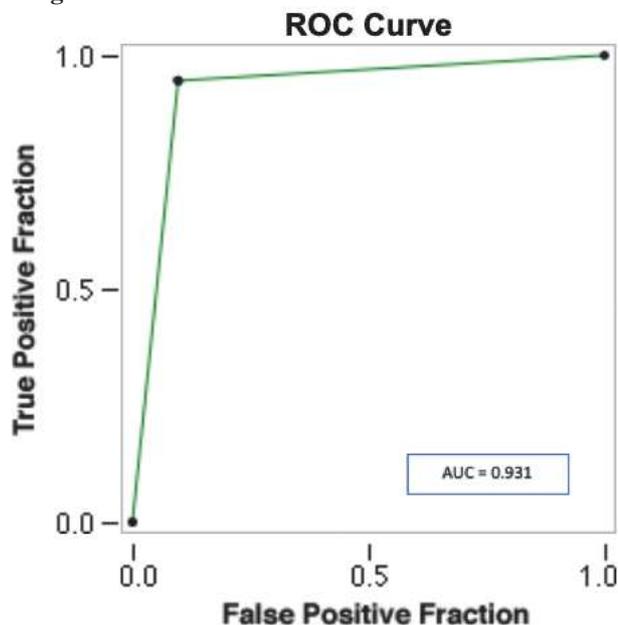
e 93% para os dataset sem e com o color deconvolution, respectivamente. Já o F-1 Score, que é a média harmônica entre precisão e *Recall* e que permite uma visão mais exata da eficiência do classificador, apresentou média de de 72% e 93% para os dataset sem e com o color deconvolution, respectivamente.

Tendo em vista que a arquitetura proposta não tenha realizado ainda nenhuma técnica de *Data Augmentation* e possui poucas camadas convolucionais, em comparação com os trabalhos relacionados, o desempenho da rede demonstrou ser promissor para lidar com o problema abordado nessa pesquisa e apresentou, dentro do contexto estudado, um desempenho considerável levando em consideração as análises realizadas com a proteína BGN.

Levando em consideração que o problema diz respeito a identificação de amostras histológicas que contenham câncer, dentro dos valores obtidos na matriz de confusão gerada pelo modelo treinado, o indicador mais relevante é o denominado verdadeiro-positivo, ou seja, quando a verificação é feita corretamente. Vale dizer também que o cenário mais perigoso é o falso-negativo, uma vez que o paciente teria a doença e o modelo o classificaria como saudável. Neste contexto, as imagens do dataset que receberam o processo de deconvolução de cores, apresetaram melhor resultado.

Realizou-se o experimento de classificação binária, conforme resumido na Figura 12, em imagens de histologia de mama saudável e com câncer. O classificador foi capaz de diferenciar 90,62% as imagens com câncer e 94,44% das imagens saudáveis corretamente. É possível observar na Figura 13 a relação entre a fração positiva verdadeira e a fração falsa positiva ilustrada na curva ROC obtida. O valor Área sob curva (AUC) obtida foi de 0.931 e intervalo de confiança de 95%: 0,882 - 0,9794.

Uma vez que o melhor resultado obtido foi através do dataset com imagens que receberam

Figura 13: Curva ROC e AUC do modelo BGNDL

Fonte: Elaborado pelo autor

o processo de deconvolução de cores, foi realizado também o teste com outra rede neural conhecida, a ResNet50, que é uma rede neural convolucional com 50 camadas de profundidade e que é também muito usada para problemas de reconhecimento de imagens. Utilizou-se a implementação por meio da biblioteca Keras Applications que são modelos de aprendizado profundo que são disponibilizados juntamente com pesos pré-treinados.

Figura 14: Comparação modelos BGNDL e ResNet50

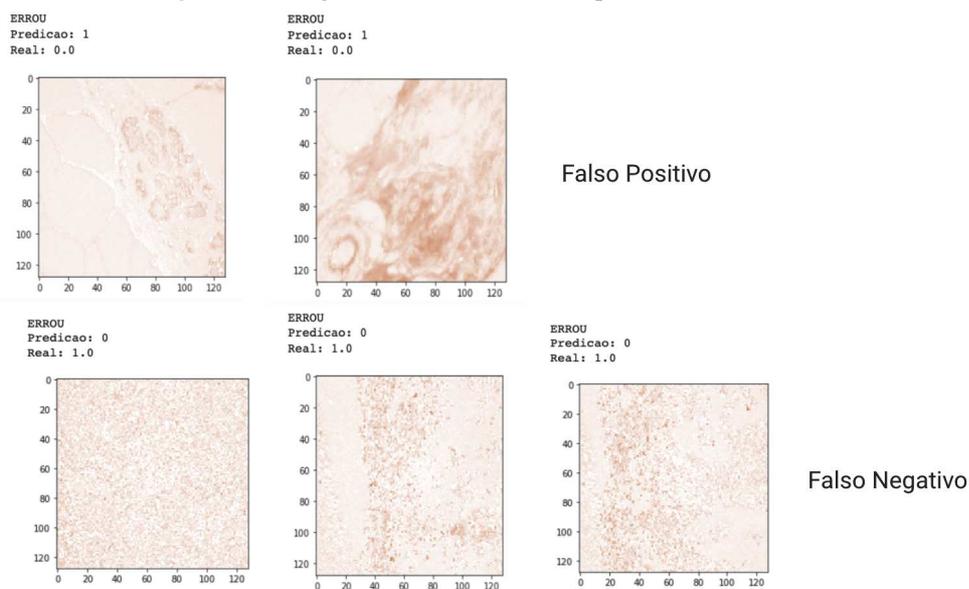
Dataset após Color Deconvolution - modelo BGNDL							
T E S T E	REAL			Acurácia	Precisão	Recall	F1-Score
	Benigno	Maligno		93%	93%	93%	93%
	Benigno	34	2				
Maligno	3	29					
Dataset após Color Deconvolution - ResNet50							
T E S T E	REAL			Acurácia	Precisão	Recall	F1-Score
	Benigno	Maligno		47%	24%	50%	32%
	Benigno	0	36				
Maligno	0	32					

Fonte: Elaborado pelo autor

Embora tenha conseguido classificar corretamente todos os casos com câncer (maligno), o modelo não acertou nenhum dos casos benignos e os classificou como malignos também. Apesar de a estrutura da rede Resnet50 ser muito mais complexa e apresentar muitas camadas convolucionais, a mesma demonstrou não ser adequada para este caso, com esse tipo específico de imagem histológica, sendo necessário mais testes para ajustes nos parâmetros.

Na Figura 15 é possível observar as imagens que foram classificadas erradas pelo modelo BGNDL. Foram classificadas duas imagens como Falso Positivo e três imagens como Falso Negativo. Apesar de não ter sido verificado nenhum padrão entre elas, é possível observar que as imagens classificadas como Falso Positivo apresentam características de maiores espaços em branco, devido ao corte histológico e as imagens Falso Negativo apresentam características de uma imagem granulada. Contudo outras imagens com essas mesmas características foram classificadas de forma correta. Deste modo, não foi possível precisar qual característica foi determinante para o modelo realizar a classificação de forma errada.

Figura 15: Imagens classificadas erradas pelo modelo BGNDL



Fonte: Elaborado pelo autor

Este trabalho apresentou a implementação de uma rede neural convolucional para classificar imagens de histologia humana. As imagens de histologia de mama foram classificadas como saudáveis e com câncer. Foi apresentada uma implementação direta, com arquitetura simples, e, até o momento sem realizar técnicas de aumento de dados.

6 CONCLUSÃO

A Inteligência Artificial trouxe resultados promissores para a Medicina. O uso de Machine Learning vem crescendo e diversas aplicações têm sido utilizadas nas áreas de detecção de doenças. Deep Learning oferece um grande conjunto de ferramentas para uso em problemas de processamento de imagens médicas. A análise patológica baseada na inspeção visual de amostras de tecido é de natureza subjetiva e existem variações entre as avaliações diagnósticas por imagem. Nesse sentido, a patologia digital, onde software e hardware dão suporte ao fluxo de trabalho do patologista, é uma grande aliada.

No contexto da imagem histológica, foco deste trabalho, nem sempre a coloração rotineira de tecidos para estudo ao microscópio é suficiente, onde nestes casos são utilizadas técnicas histoquímicas para melhorar a visualização. Uma dessas técnicas inclui o uso de marcadores biológicos, os biomarcadores. Neste trabalho buscamos verificar a diferença na expressão do biomarcador Biglycan entre tecidos mamários malignos e benignos e para isso foi criada uma rede neural convolucional.

Este trabalho apresentou a implementação de um modelo de rede neural convolucional para classificar imagens histológicas humanas, no contexto do câncer de mama, verificando a relação da proteína Biglycan e o câncer de mama como potencial biomarcador para esta doença. Para o desenvolvimento deste trabalho, foi criado um conjunto de dados original com imagens histológicas utilizando o biomarcador Biglycan. O processo de deconvolução de cores das imagens deste conjunto de dados também foi realizado, de forma automatizada, a fim de aumentar o desempenho do processo.

As imagens histológicas da mama foram classificadas por uma porcentagem média maior que 93% com o modelo BGNDL, indicando diferença entre a expressão do biomarcador Biglycan entre tecidos com e sem câncer de mama. O modelo obteve um resultado significativamente melhor com as imagens que passaram pelo processo de deconvolução de cores, com uma diferença de mais de 20% na acurácia em relação às imagens que não passaram por esse processo. Também foi realizada uma comparação com uma implementação da rede ResNet50, fornecida pelo framework Keras, que não conseguiu identificar corretamente nenhuma amostra saudável.

Embora a pesquisa tenha apresentado resultados significativos dentro da sua proposta, foi possível observar como fator limitante, dada a característica da origem dos itens do dataset, a necessidade de obter mais imagens para realizar um melhor treinamento do modelo, de forma a obter melhores resultados. Somado-se a questão da quantidade de imagens, sentiu-se também a necessidade de realizar uma avaliação da classificação das imagens que trouxeram falsos resultados, de forma a entender o que levou o modelo a realizar tal classificação, uma vez que não foi possível observar nenhum padrão que pudesse trazer a explicação desses resultados.

Com base nos resultados deste trabalho, foi feita uma contribuição à literatura, verificando a expressão do biomarcador Biglycan entre tecidos mamários saudáveis e cancerosos, utilizando redes neurais convolucionais. Para trabalhos futuros, serão utilizadas técnicas de aumento de

dados para melhorar o desempenho da rede, além da implementação para realizar a classificação de intensidade de biomarcadores. Além disso, o aprimoramento do modelo com novos possíveis algoritmos de classificação está no escopo. Como contribuição ao longo da pesquisa foi produzido um artigo para publicação em periódico, ainda em avaliação para o periódico Neurocomputing.

REFERÊNCIAS

- AEFFNER, F.; ZARELLA, M. D.; BUCHBINDER, N.; BUI, M. M.; GOODMAN, M. R.; HARTMAN, D. J.; LUJAN, G. M.; MOLANI, M. A.; PARWANI, A. V.; LILLARD, K.; TURNER, O. C.; VEMURI, V. N. P.; YUIL-VALDES, A. G.; BOWMAN, D. Introduction to Digital Image Analysis in Whole-slide Imaging: a white paper from the digital pathology association. **Journal of pathology informatics**, [S.l.], v. 10, p. 9, 3 2019.
- ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. In: INTERNATIONAL CONFERENCE ON ENGINEERING AND TECHNOLOGY (ICET), 2017., 2017. **Anais...** [S.l.: s.n.], 2017. p. 1–6.
- Albino de Queiroz, D.; André da Costa, C.; Aparecida Isquierdo Fonseca de Queiroz, E.; Folchini da Silveira, E.; da Rosa Righi, R. Internet of Things in active cancer Treatment: a systematic review. **Journal of Biomedical Informatics**, [S.l.], v. 118, p. 103814, 2021.
- ARORA, M.; TYAGI, V.; MANZAR, S. A. Computer vision based identification of abnormal tissues in biomedical images. In: INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING AND INTEGRATED NETWORKS (SPIN), 2020., 2020. **Anais...** [S.l.: s.n.], 2020. p. 37–42.
- BISCHOF, A. G.; YÜKSEL, D.; MAMMOTO, T.; MAMMOTO, A.; KRAUSE, S.; INGBER, D. E. Breast cancer normalization induced by embryonic mesenchyme is mediated by extracellular matrix biglycan. **Integrative Biology**, [S.l.], v. 5, n. 8, p. 1045–1056, 07 2013.
- Chang, J.; Yu, J.; Han, T.; Chang, H.; Park, E. A method for classifying medical images using transfer learning: a pilot study on histopathology of breast cancer. In: IEEE 19TH INTERNATIONAL CONFERENCE ON E-HEALTH NETWORKING, APPLICATIONS AND SERVICES (HEALTHCOM), 2017., 2017. **Anais...** [S.l.: s.n.], 2017. p. 1–4.
- CHANG, Y. H.; FOLEY, P.; AZIMI, V.; BORKAR, R.; LEFMAN, J. Primer for Image Informatics in Personalized Medicine. **Procedia Engineering**, [S.l.], v. 159, p. 58–65, 2016. Humanitarian Technology: Science, Systems and Global Impact 2016, HumTech2016.
- CHENG, G.; YANG, C.; YAO, X.; GUO, L.; HAN, J. When Deep Learning Meets Metric Learning: remote sensing image scene classification via learning discriminative cnns. **IEEE Transactions on Geoscience and Remote Sensing**, [S.l.], v. 56, n. 5, p. 2811–2821, 2018.
- CONG, L.; MAISHI, N.; ANNAN, D. A.; YOUNG, M. F.; MORIMOTO, H.; MORIMOTO, M.; NAM, J.-M.; HIDA, Y.; HIDA, K. Inhibition of stromal biglycan promotes normalization of the tumor microenvironment and enhances chemotherapeutic efficacy. **Breast Cancer Research**, [S.l.], v. 23, n. 1, p. 51, 2021.
- DING, B.; QIAN, H.; ZHOU, J. Activation functions and their characteristics in deep neural networks. In: CHINESE CONTROL AND DECISION CONFERENCE (CCDC), 2018., 2018. **Anais...** [S.l.: s.n.], 2018. p. 1836–1841.
- DONG, S.; WANG, P.; ABBAS, K. A survey on deep learning and its applications. **Computer Science Review**, [S.l.], v. 40, p. 100379, 2021.

- ECHLE, A.; RINDTORFF, N. T.; BRINKER, T. J.; LUEDDE, T.; PEARSON, A. T.; KATHER, J. N. Deep learning in cancer pathology: a new generation of clinical biomarkers. **British Journal of Cancer**, [S.l.], v. 124, p. 686–696, 2021.
- ERICKSON, B. J.; KORFIATIS, P.; AKKUS, Z.; KLINE, T. L. Machine Learning for Medical Imaging. **RadioGraphics**, [S.l.], v. 37, n. 2, p. 505–515, 2017. PMID: 28212054.
- 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**, [S.l.], v. 4, p. 5, 2021.
- ESTEVA, A.; KUPREL, B.; NOVOA, R. A.; KO, J.; SWETTER, S. M.; BLAU, H. M.; THRUN, S. Dermatologist-level classification of skin cancer with deep neural networks. **Nature**, [S.l.], v. 542, p. 115–118, 2017.
- FATIMA, M.; PASHA, M. Survey of Machine Learning Algorithms for Disease Diagnostic. **Journal of Intelligent Learning Systems and Applications**, [S.l.], v. Vol.09No.01, 2017.
- FUHRICH, D. G.; LESSEY, B. A.; SAVARIS, R. F. Comparison of HSCORE assessment of endometrial beta3 integrin subunit expression with digital HSCORE using computerized image analysis (ImageJ). **Analytical and quantitative cytopathology and histopathology**, [S.l.], v. 35, n. 4, p. 210–216, aug 2013.
- GIL, A. C. **Como elaborar projetos de pesquisa**. [S.l.]: Atlas, 2002.
- GUBERT, L.; COSTA, C. da; RIGHI, R. Context awareness in healthcare: a systematic literature review. **Universal Access in the Information Society**, [S.l.], 2019.
- HABUZA, T.; NAVAZ, A. N.; HASHIM, F.; ALNAJJAR, F.; ZAKI, N.; SERHANI, M. A.; STATSENKO, Y. AI applications in robotics, precision medicine, and medical image analysis: an overview and future trends. **Informatics in Medicine Unlocked**, [S.l.], p. 100596, 2021.
- HAJIAN-TILAKI, K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. **Caspian journal of internal medicine**, [S.l.], v. 4, n. 2, p. 627–635, 2013.
- HOSSEINI, M.-P.; LU, S.; KAMARAJ, K.; SLOWIKOWSKI, A.; VENKATESH, H. C. **Deep Learning Architectures** | SpringerLink. 2020.
- HUANG, M.-L.; LIN, T.-Y. Dataset of breast mammography images with masses. **Data in Brief**, [S.l.], v. 31, p. 105928, 2020.
- HUNTER, D. J.; LOSINA, E.; GUERMAZI, A.; BURSTEIN, D.; LASSERE, M. N.; KRAUS, V. A pathway and approach to biomarker validation and qualification for osteoarthritis clinical trials. **Current drug targets**, [S.l.], v. 11, n. 5, p. 536–545, may 2010.
- INCA. **O que é câncer?** Disponível em: < <https://www.inca.gov.br/o-que-e-cancer>>. Acesso em: 10 junho 2021.
- INDOLIA, S.; GOSWAMI, A. K.; MISHRA, S.; ASOPA, P. Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. **Procedia Computer Science**, [S.l.], v. 132, p. 679–688, 2018. International Conference on Computational Intelligence and Data Science.

- KAJAREE DAS, R. N. B. A Survey on Machine Learning: concept, algorithms and applications. **International Journal of Innovative Research in Computer and Communication Engineering**, [S.l.], v. 5, n. 2, p. 1301–1309, 2017.
- KATHER, J. N.; WEIS, C.-A.; MARX, A.; SCHUSTER, A. K.; SCHAD, L. R.; ZÖLLNER, F. G. New Colors for Histology: optimized bivariate color maps increase perceptual contrast in histological images. **PloS one**, [S.l.], v. 10, n. 12, p. e0145572–e0145572, dec 2015.
- KER, J.; BAI, Y.; LEE, H. Y.; RAO, J.; WANG, L. Automated brain histology classification using machine learning. **Journal of Clinical Neuroscience**, [S.l.], v. 66, p. 239–245, 2019.
- KIM, D. H.; LEE, H.; KIM, D.-H.; CHAE, S. W.; SOHN, J. H.; KIM, K.; DO, S.-I. High excision repair cross-complementation group 1 expression is associated with favorable prognostic factors in breast cancer. **Oncology Letters**, [S.l.], v. 14, n. 4, p. 4995–5003, Aug. 2017.
- KITCHENHAM, B. **Procedures for Performing Systematic Reviews**. [S.l.]: Department of Computer Science, Keele University, UK, 2004. Keele University. Technical Report TR/SE-0401.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. In: INTERNATIONAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS - VOLUME 1, 25., 2012, Red Hook, NY, USA. **Proceedings...** Curran Associates Inc., 2012. p. 1097–1105. (NIPS'12).
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, [S.l.], v. 521, n. 7553, p. 436–444, 2015.
- LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, [S.l.], v. 86, n. 11, p. 2278–2324, 1998.
- LEE, B. L.; LIEDKE, P. E.; BARRIOS, C. H.; SIMON, S. D.; FINKELSTEIN, D. M.; GOSS, P. E. Breast cancer in Brazil: present status and future goals. **The Lancet Oncology**, [S.l.], v. 13, n. 3, p. e95–e102, 2012.
- LI, B.; GE, Y.; ZHAO, Y.; GUAN, E.; YAN, W. Benign and Malignant Mammographic Image Classification Based on Convolutional Neural Networks. In: CURRAN ASSOCIATES INC., 2018, New York, NY, USA. **Anais...** Association for Computing Machinery, 2018. p. 247–251. (ICMLC 2018).
- Li, Y.; Li, X.; Xie, X.; Shen, L. Deep learning based gastric cancer identification. In: IEEE 15TH INTERNATIONAL SYMPOSIUM ON BIOMEDICAL IMAGING (ISBI 2018), 2018., 2018. **Anais...** [S.l.: s.n.], 2018. p. 182–185.
- LI, Y.; WU, J.; WU, Q. Classification of Breast Cancer Histology Images Using Multi-Size and Discriminative Patches Based on Deep Learning. **IEEE Access**, [S.l.], v. 7, p. 21400–21408, 2019.
- LITJENS, G.; KOOI, T.; BEJNORDI, B. E.; SETIO, A. A. A.; CIOMPI, F.; GHAFORIAN, M.; van der Laak, J. A.; van Ginneken, B.; SÁNCHEZ, C. I. A survey on deep learning in medical image analysis. **Medical Image Analysis**, [S.l.], v. 42, p. 60–88, 2017.

- Lu, Y.; Li, J.; Su, Y.; Liu, A. A Review of Breast Cancer Detection in Medical Images. In: IEEE VISUAL COMMUNICATIONS AND IMAGE PROCESSING (VCIP), 2018., 2018. **Anais...** [S.l.: s.n.], 2018. p. 1–4.
- MADABHUSHI, A.; LEE, G. Image analysis and machine learning in digital pathology: challenges and opportunities. **Medical Image Analysis**, [S.l.], v. 33, p. 170–175, 2016. 20th anniversary of the Medical Image Analysis journal (MedIA).
- MAHMOOD, T.; ARSALAN, M.; OWAIS, M.; LEE, M. B.; PARK, K. R. Artificial Intelligence-Based Mitosis Detection in Breast Cancer Histopathology Images Using Faster R-CNN and Deep CNNs. **Journal of Clinical Medicine**, [S.l.], v. 9, n. 3, 2020.
- MAIER, A.; SYBEN, C.; LASSER, T.; RIESS, C. A gentle introduction to deep learning in medical image processing. **Zeitschrift für Medizinische Physik**, [S.l.], v. 29, n. 2, p. 86–101, 2019. Special Issue: Deep Learning in Medical Physics.
- MESSAOUD, S.; BRADAI, A.; BUKHARI, S. H. R.; QUANG, P. T. A.; AHMED, O. B.; ATRI, M. A survey on machine learning in Internet of Things: algorithms, strategies, and applications. **Internet of Things**, [S.l.], v. 12, p. 100314, 2020.
- OLIVEIRA, R. S.-M. F. d.; OLIVEIRA, F. A. M. d.; PINHEIRO, H. S. Programando O Software Imagej para a Seleção Automática de Áreas Marcadas Por Imuno-Histoquímica e Coradas com o Cromógeno DAB e Contra-Coradas pela Hematoxilina. **Revista Interdisciplinar de Estudos Experimentais - Animais e Humanos Interdisciplinary Journal of Experimental Studies**, [S.l.], 2011.
- OLIVER, C.; JAMUR, M. C. (Ed.). **Immunocytochemical Methods and Protocols**. [S.l.]: Humana Press, 2010.
- POLO, T. C. F.; MIOT, H. A.; POLO, T. C. F.; MIOT, H. A. Aplicações da curva ROC em estudos clínicos e experimentais. **Jornal Vascular Brasileiro**, [S.l.], v. 19, p. 0–0, 12 2020.
- PONTI, M. A.; COSTA, G. B. P. **Como funciona o Deep Learning**. 2017.
- POUYANFAR, S.; SADIQ, S.; YAN, Y.; TIAN, H.; TAO, Y.; REYES, M. P.; SHYU, M.-L.; CHEN, S.-C.; IYENGAR, S. S. A Survey on Deep Learning: algorithms, techniques, and applications. **ACM Comput. Surv.**, New York, NY, USA, v. 51, n. 5, Sept. 2018.
- QUEZADA-MARÍN, J. I.; LAM, A. K.; OCHIAI, A.; ODZE, R. D.; WASHINGTON, K. M.; FUKAYAMA, M.; RUGGE, M.; KLIMSTRA, D. S.; NAGTEGAAL, I. D.; TAN, P.-H.; ARENDS, M. J.; GOLDBLUM, J. R.; CREE, I. A.; SALTO-TELLEZ, M. Gastrointestinal tissue-based molecular biomarkers: a practical categorisation based on the 2019 world health organization classification of epithelial digestive tumours. **Histopathology**, [S.l.], v. 77, n. 3, p. 340–350, 2020.
- SALVI, M.; ACHARYA, U. R.; MOLINARI, F.; MEIBURGER, K. M. The impact of pre- and post-image processing techniques on deep learning frameworks: a comprehensive review for digital pathology image analysis. **Computers in Biology and Medicine**, [S.l.], v. 128, p. 104129, 2021.
- SALVIO, A. G.; MARQUES, M. E. A. Imuno-histoquímica para identificação de células neoplásicas no infiltrado ativo de melanomas finos. **Jornal Brasileiro de Patologia e Medicina Laboratorial**, [S.l.], v. 42, p. 143–148, 4 2006.

SANTOS, R. M. M. d. **Técnicas de aprendizagem de máquina utilizadas na previsão de desempenho acadêmico**. 2016. Dissertação (Mestrado em Ciência da Computação) — Universidade Federal dos Vales do Jequitinhonha e Mucuri, 2016.

Shabbeer Basha, S. H.; Ghosh, S.; Kishan Babu, K.; Ram Dubey, S.; Pulabaigari, V.; Mukherjee, S. RCCNet: an efficient convolutional neural network for histological routine colon cancer nuclei classification. In: INTERNATIONAL CONFERENCE ON CONTROL, AUTOMATION, ROBOTICS AND VISION (ICARCV), 2018., 2018. **Anais...** [S.l.: s.n.], 2018. p. 1222–1227.

SHAO, C.; CHENG, C.; SHAO, Q.; CHEN, B. Identification and Validation of Biglycan as Prognosis and Therapy Markers for Patients with Stomach Adenocarcinoma. **International journal of general medicine**, [S.l.], v. 14, p. 3497–3509, jul 2021.

SHRESTHA, A.; MAHMOOD, A. Review of Deep Learning Algorithms and Architectures. **IEEE Access**, [S.l.], v. 7, p. 53040–53065, 2019.

SILVA, E. L. d. S.; MENEZES, E. M. **Metodologia da Pesquisa e Elaboração de Dissertação**. [S.l.]: Universidade Federal de Santa Catarina, 2001.

SRINIDHI, C. L.; CIGA, O.; MARTEL, A. L. Deep neural network models for computational histopathology: a survey. **Medical Image Analysis**, [S.l.], v. 67, p. 101813, 2021.

Srinivasan, P.; Srinivasna, V. A Comprehensive Diagnostic Tool for Skin Cancer Using a Multifaceted Computer Vision Approach. In: INTERNATIONAL CONFERENCE ON SOFT COMPUTING MACHINE INTELLIGENCE (ISCM), 2020., 2020. **Anais...** [S.l.: s.n.], 2020. p. 213–217.

THANAPOL, P.; LAVANGNANANDA, K.; BOUVRY, P.; PINEL, F.; LEPRÉVOST, F. Reducing Overfitting and Improving Generalization in Training Convolutional Neural Network (CNN) under Limited Sample Sizes in Image Recognition. In: INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY (INCIT), 2020., 2020. **Anais...** [S.l.: s.n.], 2020. p. 300–305.

T.K., B.; ANNAVARAPU, C. S. R.; BABLANI, A. Machine learning algorithms for social media analysis: a survey. **Computer Science Review**, [S.l.], v. 40, p. 100395, 2021.

TURKKI, R.; BYCKHOV, D.; LUNDIN, M.; ISOLA, J.; NORDLING, S.; KOVANEN, P. E.; VERRILL, C.; SMITTEN, K. von; JOENSUU, H.; LUNDIN, J.; LINDER, N. Breast cancer outcome prediction with tumour tissue images and machine learning. **Breast Cancer Research and Treatment**, [S.l.], v. 177, n. 1, p. 41–52, 2019.

WADHWA, S.; EMBREE, M. C.; BI, Y.; YOUNG, M. F. Regulation, Regulatory Activities, and Function of Biglycan. **Critical Reviews in Eukaryotic Gene Expression**, [S.l.], v. 14, n. 4, p. 301–316, 2004.

WATAYA, T.; NAKANISHI, K.; SUZUKI, Y.; KIDO, S.; TOMIYAMA, N. Introduction to deep learning: minimum essence required to launch a research. **Japanese Journal of Radiology**, [S.l.], v. 38, n. 10, p. 907–921, 2020.

WILEY, V.; LUCAS, T. Computer Vision and Image Processing: a paper review. **International Journal of Artificial Intelligence Research**, [S.l.], v. 2, n. 1, p. 29–36, 2018.

YANG, Q.; LIU, H.-Y.; LIU, D.; SONG, Y.-Q. Ultrasonographic Features of Triple-Negative Breast Cancer: a comparison with other breast cancer subtypes. **Asian Pacific Journal of Cancer Prevention**, [S.l.], v. 16, n. 8, p. 3229–3232, Apr. 2015.

YEH, J.-Y.; CHAN, S. CNN-Based CAD for Breast Cancer Classification in Digital Breast Tomosynthesis. In: INTERNATIONAL CONFERENCE ON GRAPHICS AND SIGNAL PROCESSING, 2., 2018, New York, NY, USA. **Proceedings...** Association for Computing Machinery, 2018. p. 26–30. (ICGSP'18).

YUNCHAO, G.; JIAYAO, Y. Application of Computer Vision and Deep Learning in Breast Cancer Assisted Diagnosis. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND SOFT COMPUTING, 3., 2019, New York, NY, USA. **Proceedings...** Association for Computing Machinery, 2019. p. 186–191. (ICMLSC 2019).

ZEISER, F. A.; da Costa, C. A.; OLIVEIRA RAMOS, G. de; BOHN, H. C.; SANTOS, I.; ROEHE, A. V. DeepBatch: a hybrid deep learning model for interpretable diagnosis of breast cancer in whole-slide images. **Expert Systems with Applications**, [S.l.], v. 185, p. 115586, 2021.

ZHAO, S.; YIN, X.; ZHAO, W.; LIU, L.; WANG, Z. Biglycan as a potential diagnostic and prognostic biomarker in multiple human cancers. **Oncology Letters**, [S.l.], Jan. 2020.