

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Paulo Vitor Costa Silva

**Classificação de imagens térmicas para
detecção do câncer de mama utilizando um
comitê de classificadores**

Uberlândia, Brasil

2023

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Paulo Vitor Costa Silva

**Classificação de imagens térmicas para detecção do
câncer de mama utilizando um comitê de classificadores**

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, como parte dos requisitos exigidos para a obtenção título de Bacharel em Ciência da Computação.

Orientador: Henrique Coelho Fernandes

Universidade Federal de Uberlândia – UFU

Faculdade de Computação

Bacharelado em Ciência da Computação

Uberlândia, Brasil

2023

Paulo Vitor Costa Silva

Classificação de imagens térmicas para detecção do câncer de mama utilizando um comitê de classificadores

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, como parte dos requisitos exigidos para a obtenção título de Bacharel em Ciência da Computação.

Trabalho aprovado. Uberlândia, Brasil, 01 de novembro de 2023:

Henrique Coelho Fernandes
Orientador

Ivan da Silva Sendin

Marcelo Zanchetta do Nascimento

Uberlândia, Brasil
2023

Agradecimentos

Agradeço primeiramente à minha família pelo apoio incondicional, compreensão e incentivo durante todo o período de estudo. Em especial, dedico uma agradecimento à minha mãe, Iris, que, apesar das adversidades, sempre se esforçou para proporcionar-me uma boa educação, oferecendo apoio no que estava ao seu alcance. Não posso deixar de mencionar minha parceira e esposa, Marcella, cujo apoio psicológico e motivação foram fundamentais para que eu me mantivesse firme em minha caminhada e não desistisse em momentos desafiadores.

Expresso minha sincera gratidão ao meu orientador Henrique, cuja orientação e suporte foram constantes ao longo de praticamente toda a minha jornada acadêmica. Sua experiência, otimismo e paciência foram essenciais para o desenvolvimento de meus projetos, começando por uma iniciação científica e concluindo-se com este projeto.

Aos amigos e colegas que compartilharam ideias, experiências e proporcionaram um ambiente colaborativo, agradeço pela troca enriquecedora que contribuiu para que eu chegasse ao final da graduação e ao amadurecimento deste trabalho. Além disso, agradeço profundamente aos meus colegas Pedro e Tiago, cujo apoio foi essencial para superar as dificuldades encontradas ao longo da graduação e pelos momentos de alegria e risadas compartilhados, tornando essa trajetória ainda mais especial.

Resumo

O câncer de mama é uma enfermidade que afeta milhares de mulheres anualmente no Brasil. A detecção precoce é crucial para aumentar as chances de recuperação, e a mamografia é o exame mais comumente utilizado para esse fim, embora possa ser desconfortável e, em alguns casos, não seja recomendada. Nesse contexto, a termografia infravermelha surge como uma alternativa atrativa para auxiliar no diagnóstico do câncer de mama. Este projeto abordou a classificação de imagens infravermelhas da mama de pacientes saudáveis e doentes, provenientes da base de dados *Database For Mastology Research* (DMR-IR), com o objetivo de identificar aquelas que apresentam a doença. Para realizar a classificação, foi feito o uso de redes neurais convolucionais pré-treinadas, bem como dois comitês de classificadores, destacando-se o comitê do tipo voto majoritário. Este último apresentou resultados notáveis quando comparado com cada CNN individualmente, atingindo uma acurácia de 90,91% e um F1-Score de 92,31%. Esses resultados validam a eficácia do uso de comitês de classificadores como uma estratégia promissora para aprimorar a detecção de câncer de mama por meio de imagens infravermelhas.

Palavras-chave: câncer, imagens, termografia, classificação, comitê.

Lista de ilustrações

Figura 1 – Mamografias e diferentes graus de densidade do tecido da mama - extraído de Society (2023)	10
Figura 2 – Exemplo de imagem contida na base de dados DMR-IR, colorida com filtro da biblioteca matplolib.	12
Figura 3 – O que é o câncer? - extraído de INCA (2023)	16
Figura 4 – Espectro eletromagnético - traduzido de Ibarra-Castanedo (2005). . . .	18
Figura 5 – Rede Neural Convolutacional e suas diferentes camadas - extraído de Chaves (2019).	19
Figura 6 – Exemplificação da construção de um Comitê de Classificadores.	20
Figura 7 – Fluxograma das etapas abordadas	24
Figura 8 – Gráfico gerado após treinamento em um dos experimentos com a rede VGG16	26
Figura 9 – Gráfico gerado após treinamento em um dos experimentos com a rede VGG16, a critério de comparação	26

Lista de tabelas

Tabela 1 – Média dos resultados obtidos em 10 execuções para cada modelo	27
Tabela 2 – Resultados obtidos por cada comitê de classificadores	27

Lista de abreviaturas e siglas

CAD	Sistemas de diagnóstico auxiliado por computador
CNN	Redes neurais convolucionais (Convolutional Neural Networks)
DL	Aprendizado Profundo (Deep Learning)
DMR-IR	Database for Mastology Research
FP	Falso Positivo
FN	Falso Negativo
HUAP	Hospital Universitário Antônio Pedro
SVM	Maquina de Vetores de Suporte (Support Vector Machine)
TIR	Termografia Infravermelha
VP	Verdadeiro Positivo
VN	Verdadeiro Negativo

Sumário

1	INTRODUÇÃO	9
1.1	Justificativa	11
1.2	Objetivos	11
1.3	Métodos	12
1.3.1	Base de dados	12
1.3.2	Redes pré-treinadas	13
1.3.3	Comitê de classificadores	13
1.3.4	Métricas de avaliação	14
1.4	Organização deste trabalho	15
2	REFERENCIAL TEÓRICO	16
2.1	Conceitos Fundamentais	16
2.1.1	Câncer de Mama	16
2.1.2	Termografia Infravermelha (TIR)	17
2.1.3	Redes Neurais Convolucionais	18
2.1.4	Comitê de Classificadores	19
2.2	Trabalhos Correlatos	21
3	DESENVOLVIMENTO	24
3.1	Ajuste na base de dados	25
3.2	Aumento de dados	25
3.3	Classificação com as redes pré-treinadas	25
4	RESULTADOS E DISCUSSÃO	27
5	CONCLUSÃO	30
	REFERÊNCIAS	32

1 Introdução

Segundo os dados mais recentes do Instituto Nacional do Câncer, estima-se que o câncer de mama é o maior responsável pelos novos casos de câncer em mulheres no Brasil, com uma taxa de 30,1% no ano de 2022. Já no ano de 2020, a estimativa de mortalidade conforme a localização do tumor, aponta uma taxa de 16,5% de óbitos por câncer de mama, que também é a maior comparada aos outros tipos de câncer (INCA, 2023).

O câncer de mama, em grande parte dos casos, pode ser detectado nos estágios iniciais da doença, tornando seu tratamento bastante eficaz, com altas taxas de sucesso. Logo, a detecção precoce da doença é fundamental para aumentar as taxas de efetividade do tratamento e diminuir o índice de mortalidade decorrente da doença. O exame mais comum utilizado atualmente para a detecção do câncer de mama, e recomendado pelo Ministério da Saúde, é a mamografia por rastreio, capaz de identificar alterações suspeitas de câncer antes do surgimento dos sintomas, ou seja, antes que seja palpada qualquer alteração nas mamas (INCA, 2023). Esse exame é oferecido para mulheres com mais de 40 anos no Brasil, pelo Sistema Único de Saúde (SUS).

A mamografia por rastreio apresenta bastante pontos negativos, como o alto índice de resultados incorretos, em sua maioria falsos positivos, além de submeter a paciente a exposição aos raios X durante a realização do exame (INCA, 2023). Existe também como opção, a mamografia diagnóstica, que pode ser solicitada por um médico para pacientes de qualquer idade. Essa abordagem não é eficiente no caso de mulheres jovens, visto que as mamas de pacientes jovens são compostas principalmente por tecido glandular, ou seja, são mamas mais densas o que torna a visualização e identificação do câncer um desafio (BORCHARTT, 2013).

Na Figura 1, são apresentadas quatro imagens de mamografias, extraídas de Society (2023) onde são representadas as quatro categorias de densidade da mama, sendo a mama mais à esquerda majoritariamente composta por tecido adiposo (mais comum em mulheres mais velhas) e a mais à direita é extremamente densa (mais comum em mulheres mais jovens).

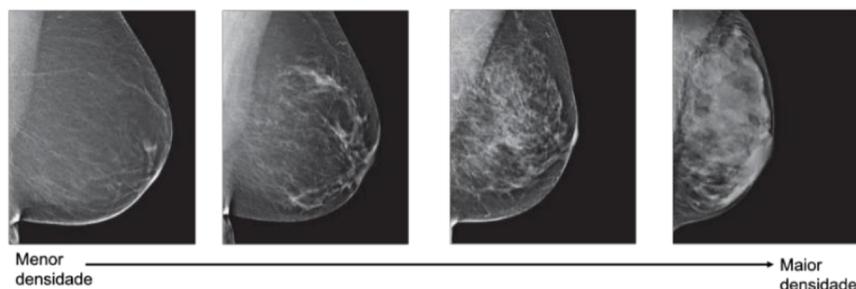


Figura 1 – Mamografias e diferentes graus de densidade do tecido da mama - extraído de [Society \(2023\)](#)

Geralmente, os primeiros tipos de procedimento realizados em busca de alterações suspeitas na mama, são o autoexame e o exame clínico. Ambos são procedimentos manuais no qual a paciente ou um especialista, apalpa a mama na busca por alterações ([BARCELOS, 2017](#)). Existem diversos exames que utilizam imagens como método para a detecção do câncer de mama e que apresentam altos índices informativos relevantes para o profissional que está realizando a análise. No entanto, em grande maioria possuem como desvantagens altos custos para a realização do exame ([BORCHARTT, 2013](#)).

Dentre as técnicas de suporte ao diagnóstico que utilizam imagens, a Termografia Infravermelha (TIR) é uma técnica não invasiva, de baixo custo e que não envolve radiação nociva ao ser humano. Seu funcionamento consiste em registrar a radiação emitida pela superfície da pele da paciente, através de uma câmera térmica, afim de determinar a sua temperatura ([BARCELOS, 2017](#)). As variações de temperatura do tecido canceroso em relação a um tecido saudável se dão pelo processo de angiogênese ([AZEVEDO et al., 2018](#)), processo pelo qual as células se proliferam descontroladamente por causa de mutações genéticas, passando a se reproduzir exageradamente, o que exige maior quantidade de nutrientes no local, gerando assim, novos vasos sanguíneos ao redor das células ([BORCHARTT, 2013](#)). Desse modo, caso exista uma anomalia no tecido da mama, haverá uma alteração de temperatura proveniente do processo de angiogênese, acarretando em regiões mais quentes, possibilitando a identificação através do uso de câmeras térmicas ([BARCELOS, 2017](#)).

As câmeras termográficas captam a radiação térmica emitida pelo corpo e a convertem em uma imagem que representa a distribuição de temperaturas superficiais. Várias técnicas e métodos computacionais tem sido propostos e aplicados afim de salientar as informações contidas nessas imagens. Desse modo, para a análise e classificação dos dados, faz-se o uso de técnicas de processamento de imagens, com suporte das áreas de aprendizado de máquina e reconhecimento de padrões ([CARDOSO, 2018](#)).

Existem muitos trabalhos relacionados ao uso da TIR para detecção do câncer de mama, e com o avanço tecnológico nos equipamentos de câmeras infravermelhas após o

ano 2000, vários pesquisadores consideram o uso desta técnica de imageamento bastante promissora. Logo, o uso da TIR é uma atraente técnica para facilitar a detecção precoce do câncer de mama (BARCELOS, 2017).

1.1 Justificativa

Os Sistemas de Detecção ou Diagnóstico Auxiliado por Computador (CAD) estão se tornando cada vez mais comuns, e frequentemente novas técnicas, como a termografia, estão sendo integradas a esses sistemas. Um sistema CAD é uma ferramenta computacional ou algoritmo capaz de identificar padrões em diferentes tipos de dados (ZULUAGA-GOMEZ et al., 2021). O rápido crescimento da colaboração virtual, das ferramentas de programação e do desempenho computacional despertou o interesse de muitos pesquisadores na aplicação de sistemas CAD na área biomédica. Como resultado, na literatura, diversos estudos têm investigado e proposto CADs para auxiliar no diagnóstico de câncer de mama (ZULUAGA-GOMEZ et al., 2021).

Com o crescente uso de sistemas de classificadores em diversas tarefas, também aumentou-se a busca por sistemas cada vez mais eficientes. Um dos interesses crescentes abordados em alguns trabalhos, trata-se da combinação ou integração de classificadores, visando melhorias de desempenho. Sistemas que envolvem a fusão de dois ou mais classificadores são conhecidos como comitês de classificadores, podendo ser compostos por classificadores do mesmo tipo ou de tipos diferentes, utilizando algoritmos de aprendizado de máquina distintos (SILVA, 2017).

Diversas pesquisas recentes abordam imagens termográficas da mama, evidenciando a relevância desse tema na comunidade científica. A análise das imagens termográficas da mama tem sido conduzida por diversas perspectivas, incluindo a busca por aprimoramentos na segmentação e a aplicação de técnicas de extração de características para fins de classificação. Apesar da diversidade de abordagens técnicas nestes estudos, a exploração de comitês de classificadores foi subutilizada na melhoria do diagnóstico, embora estes tenham demonstrado resultados promissores em sua aplicação (CARDOSO, 2018; CHAVES, 2019; ZULUAGA-GOMEZ et al., 2021).

1.2 Objetivos

Este investiga o treinamento de um comitê de classificadores baseado em redes neurais convolucionais (CNNs), separando pacientes saudáveis e pacientes com câncer. Para isso, foram utilizadas imagens termográficas da mama. Os objetivos específicos deste trabalho de graduação foram:

- compreender os fundamentos envolvidos na termografia infravermelha, aprendizagem profunda (DL) e CNNs;
- utilizar redes pré-treinadas para desenvolver os classificadores;
- combinar os resultados individuais das redes para criar um comitê de classificadores;
- comparar a eficácia deste modelo, através de métricas como acurácia e F1-score, de abordagens estáticas já desenvolvidas em projetos passados;
- comparar os resultados obtidos com outros trabalhos da literatura.

1.3 Métodos

Para a execução deste trabalho serão utilizados alguns métodos fundamentais, e a compreensão básica dos mesmos é necessária. Assim, alguns destes métodos são melhores detalhados a seguir.

1.3.1 Base de dados

Em [Silva et al. \(2014\)](#) é apresentada a base de dados DMR-IR (Database for Mastology Research), a qual é composta por imagens termográficas da mama, comumente utilizada na literatura, e que foi adotada também neste projeto. Originalmente, a base era composta de dados de 141 pacientes, coletados no Hospital Universitário Antônio Pedro (HUAP) da Universidade Federal Fluminense, entretanto, a base contou com a adição de imagens de novos pacientes ao longo do tempo.

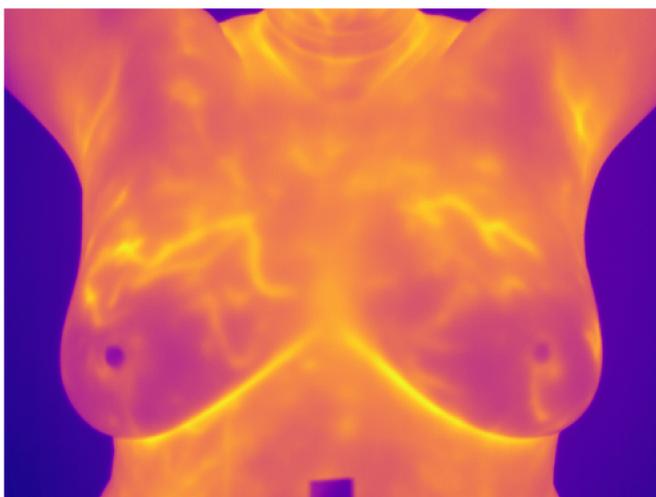


Figura 2 – Exemplo de imagem contida na base de dados DMR-IR, colorida com filtro da biblioteca matplotlib.

Cada paciente tem coletadas imagens tanto com protocolo estático quanto com o protocolo dinâmico. Cada paciente possui 5 imagens coletadas em diferentes posições utilizando o protocolo estático, sendo elas: frontal, laterais direita e esquerda (45°) e laterais direita e esquerda (90°). Para o protocolo dinâmico, foi proposto pelos autores uma coleta na qual aclimata-se a paciente até que a média de temperatura da região entre as mamas atinja $30^\circ C$ ou que o tempo máximo tenha sido atingido (5 minutos). Após a aclimação da paciente, foram coletadas 20 imagens em um período de 5 minutos e ainda 2 outras imagens laterais direita e esquerda de (90°). Neste projeto, serão utilizadas as imagens de protocolo estático, com tamanho de 640×480 pixels. As imagens se encontram publicamente em [Visual Lab DMR](#), e a base possui as devidas autorizações dos comitês de ética competentes.

1.3.2 Redes pré-treinadas

A primeira etapa do trabalho consiste na seleção de algumas CNNs utilizadas na literatura como critério de comparação, fazendo o uso de técnicas de transferência de aprendizado. Diversas redes disponíveis na literatura foram treinadas com a base ImageNet, a qual contém milhares de registros, possuindo bons extratores de características que poderiam ser reutilizados para novos problemas. Nesse contexto, foram selecionadas três redes para o estudo, tendo como critério o uso comum entre os trabalhos que utilizam imagens termográficas e que obtiveram bons resultados para o problema em questão. As redes selecionadas foram VGG16, RESNET50 E DENSENET201.

O trabalho será desenvolvido utilizando o [PyTorch](#), biblioteca de aprendizado de máquina de código aberto baseada na biblioteca Torch, usada em aplicativos de visão computacional e processamento de linguagem natural. Os classificadores serão utilizados fazendo o uso da transferência de aprendizado, uma vez que existem poucas imagens no banco e um treinamento completo não é possível, mantendo somente a camada densamente conectada e mantendo as camadas convolucionais que atuam como extratores de características das redes originais.

1.3.3 Comitê de classificadores

A etapa subsequente do estudo, consiste na coleta dos resultados individuais das redes neurais pré-treinadas. O comitê de classificadores será construído combinando os resultados obtidos individualmente, visando um desempenho aprimorado no diagnóstico. Existem diversos modelos de comitês, explorados em seções posteriores deste trabalho.

1.3.4 Métricas de avaliação

Existem diversas métricas utilizadas para a avaliação dos classificadores, sendo as mais comuns e que serão utilizadas neste projeto: acurácia, sensibilidade, especificidade e F1-Score. Para que seja possível entender o significado dessas métricas, é necessário que os conceitos por trás de seus cálculos sejam definidos, sendo eles de: verdadeiro positivo (VP), falso positivo (FP), verdadeiro negativo (VN) e falso negativo (FN). As definições utilizadas no contexto das imagens médicas são (BARATLOO et al., 2015):

- **Verdadeiro Positivo:** Número de amostras rotuladas como doentes e que são identificadas como doentes.
- **Falso Positivo:** Número de amostras rotuladas como saudáveis e que são identificadas como doentes.
- **Verdadeiro Negativo:** Número de amostras rotuladas como saudáveis e que são identificadas como saudáveis.
- **Falso Negativo:** Número de amostras rotuladas como doentes e que são identificadas como saudáveis.

Definidos estes conceitos, as métricas citadas podem ser definidas como segue:

- **Acurácia:** Capacidade do modelo diferenciar a imagem nas classes saudável e doente de forma correta, e é calculada a partir das proporções de VP's e VN's sobre todos os casos avaliados.

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (1.1)$$

- **Sensibilidade:** Capacidade do modelo determinar as imagens rotuladas como doente de forma correta.

$$Sensibilidade = \frac{VP}{VP + FN} \quad (1.2)$$

- **Especificidade:** Capacidade do modelo determinar as imagens rotuladas como saudáveis de forma correta.

$$Especificidade = \frac{VN}{VN + FP} \quad (1.3)$$

- **F1-Score:** É a média harmônica entre precisão e recall. Esta é uma métrica muito eficiente quando se possui um dataset (conjunto de dados) com classes desproporcionais, e seu modelo não emite probabilidades.

$$F1\text{-Score} = \frac{VP}{VP + 0.5 \times (FP + FN)} \quad (1.4)$$

1.4 Organização deste trabalho

O presente trabalho foi dividido em cinco capítulos, sendo este o capítulo introdutório, transmitindo uma visão geral do trabalho, e os posteriores:

Capítulo 2 - Referencial Teórico: capítulo no qual são apresentados e discutidos os trabalhos relacionados presentes na literatura, descrevendo as técnicas utilizadas assim como alguns resultados obtidos pelos autores. Também são discutidos conceitos necessários para a compreensão deste trabalho, tais como, conceitos sobre o câncer de mama e como se dá o processo de carcinogênese. Também são abordados alguns dos principais exames de imageamento utilizados no prognóstico da doença, além da termografia infravermelha que é o exame proposto nesse trabalho.

Capítulo 3 - Desenvolvimento: neste capítulo, aprofundou-se o processo de desenvolvimento deste trabalho, explorando mais detalhadamente os métodos propostos.

Capítulo 4 - Resultados e Discussão: o capítulo apresenta os resultados esperados e obtidos durante a execução dos experimentos, incluindo uma comparação com os resultados alcançados por outros autores em trabalhos semelhantes.

Capítulo 5 - Conclusão: por fim, são apresentadas as conclusões acerca dos resultados obtidos e o que se era esperado do trabalho.

2 Referencial Teórico

2.1 Conceitos Fundamentais

Nesta seção, são abordados conceitos fundamentais necessários para o entendimento do projeto.

2.1.1 Câncer de Mama

Segundo [INCA \(2023\)](#), o termo câncer abrange mais de 100 diferentes tipos de doenças malignas e benignas que possuem em comum o crescimento desordenado das células, dividindo-se rapidamente e provocando a formação de tumores, podendo invadir tecidos adjacentes, afetar diferentes órgãos e que podem se espalhar para outras regiões do corpo através de um processo conhecido como metástase.

Os diferentes tipos de câncer podem ser diferenciados pelo processo da metástase e também são correspondentes aos vários tipos de células do corpo, sendo denominadas **carcinomas** quando começam em tecidos epiteliais como a pele ou mucosas. Se o ponto de início são tecidos conjuntivos, como ossos, músculo ou cartilagem, então o câncer é classificado como um **sarcoma** ([INCA, 2023](#)).

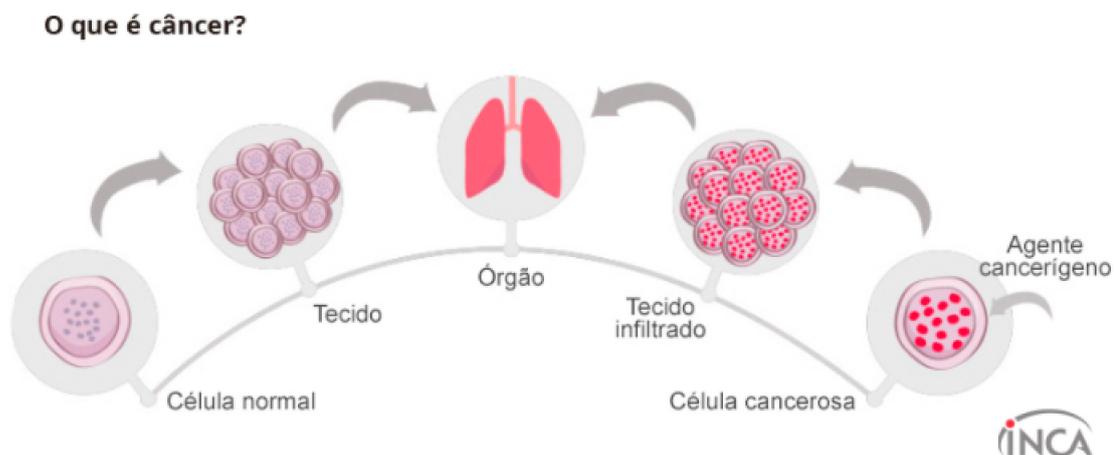


Figura 3 – O que é o câncer? - extraído de [INCA \(2023\)](#)

A Figura 3 ilustra a diferença entre o tecido que contém células normais e o tecido composto por células contendo o agente cancerígeno, exemplificando o que é o câncer. O câncer de mama está na classe dos carcinomas, e é uma das principais doenças que podem acometer a mama, juntamente com o fibroadenoma e as alterações fibrocísticas ([BARCELOS, 2017](#)). Existem vários tipos de câncer de mama, nos quais alguns possuem

um rápido desenvolvimento, enquanto outros podem crescer lentamente (INCA, 2023). Dentre alguns dos tipos mais comuns de câncer na mama, estão os carcinoma ductal e carcinoma lobular, os quais diferem pela região de início no órgão, e também o câncer inflamatório.

Segundo pesquisas realizadas em Barcelos (2017), as alterações fibrocísticas e o fibroadenoma são alterações benignas, sendo o primeiro capaz de se manifestar pela presença de cistos, dores e secreções. Já o fibroadenoma, que é um tipo de tumor benigno bastante comum encontrado na mama, provoca o crescimento de nódulos e pode ser desenvolvido durante o período reprodutivo da mulher.

A técnica mais comum de imageamento para a detecção do câncer de mama é a mamografia, que é um procedimento padrão realizado com o uso de um mamógrafo, o qual produz uma imagem pela irradiação de feixes ionizantes (raios X) de baixa dosagem na mama da paciente. A mamografia é capaz de detectar massas e calcificações na mama que podem ser indícios de câncer. Parte dos feixes de radiação é absorvida pelo tecido, assim como em outros tipos de exames que também utilizam raios X. A absorção é proporcional a densidade da mama, logo, mamas compostas em sua maioria por tecido glandular denso, acabam absorvendo muita radiação, o que aumenta as chances de problemas decorrentes do uso da radiação e ainda dificulta a visualização das calcificações, já que o tecido denso pode mascarar ou esconder regiões cancerígenas (BARCELOS, 2017).

2.1.2 Termografia Infravermelha (TIR)

Segundo pesquisas realizadas em Forcelini et al. (2019), a TIR caracteriza-se como uma técnica de imageamento que permite o registro da distribuição da radiação térmica pela superfície de um corpo, transformando-a, por meio das leis da física, em valores de temperatura. O princípio de funcionamento da TIR se dá por meio da captura da energia infravermelha emitida pelos corpos, a qual não é visível a olho nu.

Assim, devido ao processo de angiogênese, caso haja algum tipo de anomalia no tecido da mama, haverá uma alteração de temperatura nesta região, tornando viável o uso da TIR. As imagens termográficas são obtidas através de câmeras capazes de detectar radiação infravermelha.

Para que seja possível entender melhor o funcionamento da TIR, pode-se definir o conceito de corpo negro. Um corpo negro é um objeto ideal, não existente na realidade, cuja emissividade é 1. O corpo negro é um absorvedor e radiador perfeito, isto é, é capaz de absorver e refletir toda radiação incidente sobre ele, em todos os comprimentos de onda. Sendo assim, nenhum outro objeto ou superfície é capaz de emitir mais energia que o corpo negro (FORCELINI et al., 2019). Isto significa que nem o corpo humano e nem os demais objetos são capazes de emitir a mesma quantidade de radiação que o corpo

negro emitiria nas mesmas condições. Portanto, a emissividade é dada pela relação entre a radiação emitida pelo objeto real e a que seria emitida pelo corpo negro. A TIR não é influenciada apenas pela emissividade, dado que o imageamento não ocorre no vácuo, sendo necessário considerar também a transmissividade e refletividade do próprio objeto e ainda a radiação refletida por outros objetos no ambiente e a influência da atmosfera na imagem a ser coletada (BARCELOS, 2017).

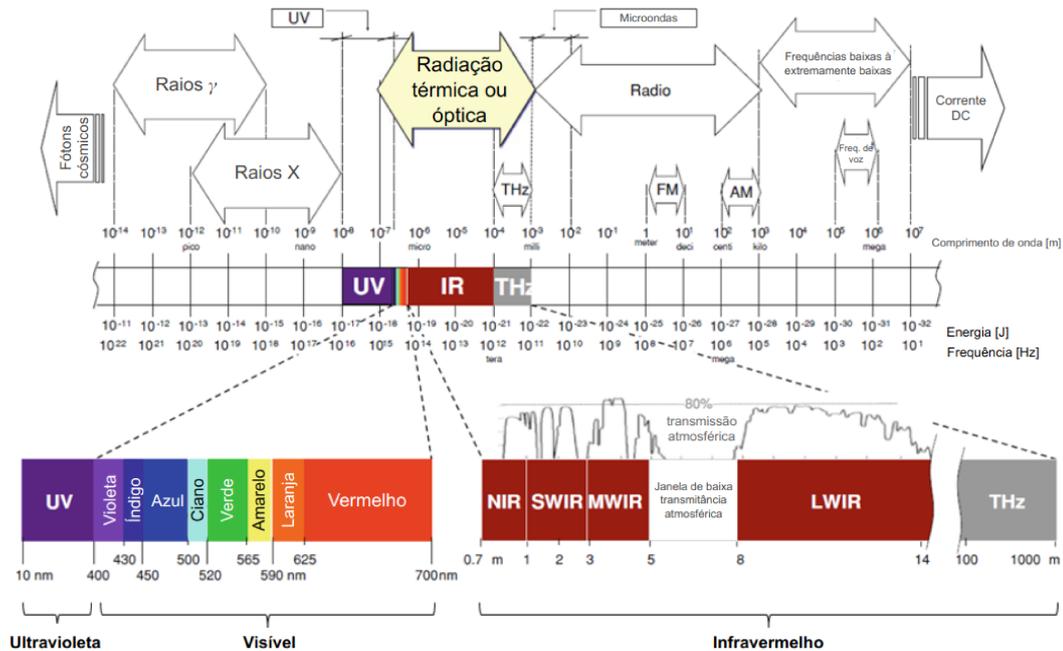


Figura 4 – Espectro eletromagnético - traduzido de Ibarra-Castanedo (2005).

A imagem da Figura 4 ilustra o espectro eletromagnético, onde suas faixas são definidas de acordo com o comprimento de onda. O espectro infravermelho abrange comprimentos de onda entre $0,74\mu\text{m}$ e $1000\mu\text{m}$. Este é um espectro que se situa entre os espectros da luz visível e das microondas, e é uma das bandas que compõem a radiação térmica (BARCELOS, 2017; BORCHARTT, 2013). A visualização no espectro infravermelho se dá através das imagens termográficas, obtidas por câmeras termográficas.

2.1.3 Redes Neurais Convolucionais

Uma rede neural convolucional ou CNN (do inglês, Convolutional Neural Networks), é uma arquitetura de rede neural artificial, variante do perceptron de múltiplas camadas (multicamadas) baseadas em um processo biológico: processamento de dados visuais. As CNNs são comumente utilizadas em aplicações cujo o objetivo consiste na detecção, classificação e reconhecimento de imagens e vídeos, e vem apresentando resultados promissores (CHAVES, 2019). As CNNs são divididas em alguns estágios principais, os quais podem ser vistos na Figura 5.

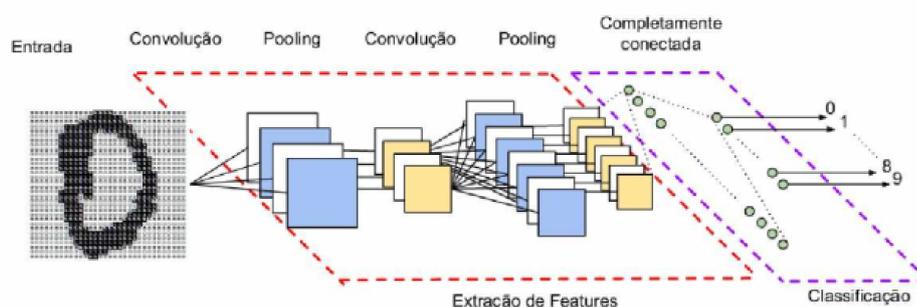


Figura 5 – Rede Neural Convolutional e suas diferentes camadas - extraído de [Chaves \(2019\)](#).

Os dois principais tipos de camadas que constituem os primeiros estágios de uma CNN, são as camadas de convolução e as camadas de pooling, enquanto que os outros estágios são compostos pelas camadas completamente conectadas. As camadas de convolução são responsáveis pela extração de características das imagens de entrada, ou seja, geram o que é chamado de *feature map*. Nessas camadas, os neurônios são responsáveis pela aplicação de filtros, assim como em outros processos mais tradicionais de visão computacional, e é comum que em cada camada de convolução sejam utilizados mais de um filtro, ou seja, existem mais de um neurônio que conecta uma mesma região da imagem de entrada ([CHAVES, 2019](#)).

Após a camada de convolução, é comum que as CNNs do estado da arte utilizem *Batch Normalization* (BN) para efetuar a normalização dos mapas de características (*feature maps*), de forma que a ativação média se mantenha próxima à 0 e o desvio padrão próximo à 1 ([CHAVES, 2019](#)).

As camadas de pooling, que são criadas após as camadas de convolução, tem como objetivo a diminuição da dimensão espacial dos dados de entrada, possibilitando que as camadas mais profundas possam ter mais filtros sem que haja o aumento do custo computacional por camada, além de evitar o *overfitting* ([CHAVES, 2019](#)).

A última camada é a camada completamente conectada, a qual tem como objetivo classificar as características extraídas pelas camadas de convolução e pooling, em alguma classe pré-definida ([CHAVES, 2019](#)).

2.1.4 Comitê de Classificadores

O uso de comitês, ou *máquinas de comitê* no aprendizado de máquina é crescente há alguns anos. É uma técnica de classificação que consiste na união de diversos classificadores, considerada bastante eficiente se comparada a utilização de apenas um classificador, pois proporciona maior confiabilidade ([VOGADO et al., 2017](#)).

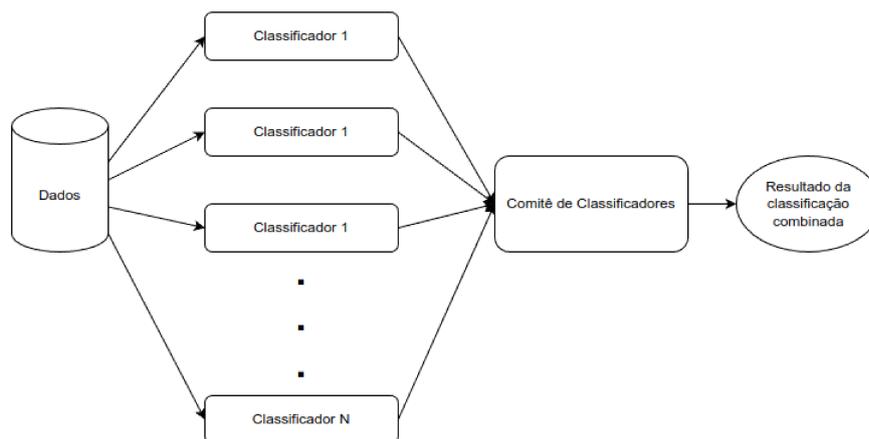


Figura 6 – Exemplificação da construção de um Comitê de Classificadores.

Como pode ser visto na Figura 6, o comitê utiliza a combinação de resultados individuais das redes neurais para conseguir um melhor desempenho de generalização. Os comitês podem ser divididos em duas grandes categorias, sendo elas as estruturas estáticas e as estruturas dinâmicas (HAYKIN, 1998 apud SILVA, 2005).

Nas estruturas estáticas, as saídas das redes neurais são combinadas sem envolver o sinal de entrada das redes. Esta categoria possui dois métodos de aplicação do sinal de entrada, que são a média de ensemble e o reforço (SILVA, 2005).

No método média de ensemble, utiliza-se o mesmo sinal de entrada para todas as redes e a saída global é dada pela combinação linear das saídas de cada rede. Já o método de reforço é caracterizado pela diferente distribuição do sinal de entrada das redes (HAYKIN, 1998 apud SILVA, 2005).

Nos comitês de classificadores com estruturas dinâmicas, o sinal de entrada está aplicado não apenas as entradas das redes, mas também sobre o combinador das saídas das redes. Desse modo, a saída do comitê depende do sinal de entrada e da saída das redes. Dentre suas técnicas, a principal é a mistura de especialistas (redes neurais), que apresenta como característica uma unidade denominada de rede de passagem, responsável pela combinação das saídas das redes (HAYKIN, 1998 apud SILVA, 2005).

Segundo Silva (2017), os comitês de classificadores podem ser divididos em multiclassificadores homogêneos e multiclassificadores heterogêneos. Dentre os multiclassificadores homogêneos, destacam-se os comitês do tipo *Bagging* e *Boosting* que trabalham utilizando amostragens do conjunto de dados original como dados de treinamento e combinam os resultados obtidos com os diferentes classificadores base através de algum esquema de votação.

Os multiclassificadores heterogêneos, utilizam diferentes algoritmos de aprendizado de máquina, tendo como dados de treinamento o mesmo conjunto de dados original. Dentre

os tipos de comitês desta modalidade, os mais utilizados são *Voting* e *Stacking* (SILVA, 2017), ambos utilizados para elaboração deste trabalho.

Um comitê do tipo *Voting*, requer que cada classificador base emita um voto para a classe que ele considera mais provável para aquela determinada instância do conjunto de dados. Em seguida, é feita uma contagem do número de votos para cada classe, proveniente dos classificadores. A classe que obtiver o maior número de votos é selecionada como a predição final para a instância de dado. Este esquema de votação, é um dos esquemas mais simples de ser executado, e é conhecido como votação majoritária (SILVA, 2017).

Já um comitê do tipo *Stacking*, espera a saída dos classificadores base como entrada para um meta-classificador, que combina essas saídas em uma predição final, após seu treinamento (SILVA, 2017).

2.2 Trabalhos Correlatos

Diversos trabalhos utilizam imagens termográficas da mama para a identificação da presença de tumores malignos e benignos ou algum tipo de alteração na mama. Esses trabalhos diferenciam-se pela forma na qual são extraídas características e o classificador utilizado, sendo mais comum pesquisas que utilizam imagens estáticas.

Em Andrade, Paiva e Correa (2017), é proposta uma metodologia para a transformação de sinais térmicos em séries temporais, onde são extraídas características para a tarefa de classificação. As imagens utilizadas na análise seguem o protocolo dinâmico e pertencem à base de dados Database for Mastology Research with Infrared Image (DMR-IR), adquiridas no Hospital Antônio Pedro da Universidade Federal Fluminense (HUAP-UFF). Neste trabalho, foi realizado o procedimento de pré-processamento afim de converter os termogramas capturados em um formato que facilitasse a análise das informações relevantes ao problema.

As etapas de pré-processamento, construção da série temporal e extração de características não possuem informações sobre o uso de algum software externo, o que sugere que foram realizadas de forma manual. Durante a etapa de classificação, fez-se o uso de alguns classificadores, sendo eles a Máquina de Vetores de Suporte (SVM) e K-Star. Para a seleção de características, também fez-se o uso da ferramenta Weka com o método de busca BeFirst e avaliação de atributos SubsetEval. Os primeiros resultados apresentados fazendo os testes com as séries originais, não apresentando resultados satisfatórios e, para contornar o problema, foi adicionada mais uma etapa de re-amostragem, dividindo a série original em janelas de 2 tempos. Com isso, os resultados obtidos pelo classificador K-Star foram uma acurácia de 95,8%, sensibilidade 93,6% e especificidade 95,9%. Já através do classificador SVM, os resultados apresentados foram uma acurácia de 91%, sensibilidade de 85% e especificidade 91%.

Um trabalho muito similar é realizado em [Cardoso \(2018\)](#), onde também é proposta uma metodologia para distinguir exames de pacientes em saudáveis e não saudáveis utilizando imagens de termografia infravermelha dinâmica, realizando a modelagem do problema utilizando séries temporais. As imagens utilizadas também pertencem a base de dados DMR-IR. Durante o pré-processamento é gerada uma super-série temporal para cada paciente, e para a extração de características sobre as séries construídas foi utilizado o método da Matrix Profile. As subsequências apontadas pelo método são concatenadas gerando uma nova série temporal reduzida. Na etapa de classificação, foi utilizada a biblioteca *Scikit-learn* que contém uma implementação do SVM em Python. Das métricas de avaliação, foram utilizadas a sensibilidade, precisão e acurácia, tendo como melhores resultados, 70%, 80, 17% e 75% respectivamente, sem o uso de normalizações e utilizando a metodologia considerando apenas *discords* (um dos resultados apresentados do método Matrix Profile).

O trabalho realizado em [Chaves \(2019\)](#), tem como objetivo a detecção do câncer de mama por meio de imagens infravermelhas de protocolo estático, também pertencentes a base de dados DMR-IR, com 440 imagens de 88 pacientes distintas, sendo 220 rotuladas como normal e 220 com patologia. Foram utilizadas redes neurais convolucionais pré-treinadas e com bom desempenho comprovado, classificando-as como portadoras ou não de câncer de mama. Diversas CNNs do estado da arte foram avaliadas, dentre elas alexNet, googleNet, resNet-18, vgg-16 e vgg-19. A implementação foi feita utilizando a linguagem Matlab. Como métricas de avaliação, foram utilizadas acurácia, sensibilidade e especificidade, as quais tiveram melhor resultado utilizando as CNNs vgg-16 com 77, 5% de acurácia, 85% de sensibilidade e 70% de especificidade, e vgg-19 com 77, 5% de acurácia, 90% de sensibilidade e 65% de especificidade.

Um outro trabalho, que também utiliza imagens termográficas de protocolo dinâmico encontradas na base de dados DMR-IR e CNNs para detecção do câncer de mama, é proposto em [Zuluaga-Gomez et al. \(2021\)](#). Neste são consideradas 57 pacientes contendo 20 imagens cada, sendo 19 pacientes saudáveis e 37 pacientes doentes. Para o pré-processamento são extraídas as regiões de interesse (ROIs) contendo somente a região das mamas.

Ainda no trabalho de [Zuluaga-Gomez et al. \(2021\)](#), as CNNs do estado da arte utilizadas avaliadas foram ResNet, SeResNet, vgg-16, Inception, InceptionResNetV2, Xception e também *baseline* que são CNNs propostas pelos autores. Além disso, CNNs com otimização dos hiper parâmetros foram propostas. Dentre os parâmetros considerados na otimização tem-se: número mínimo e máximo de blocos; número de camadas convolucionais e número de filtros por bloco, tipo do otimizador, tamanho kernel, tamanho da camada de pooling, tipo de batch normalization, taxa de dropout, número de neurônios conectados nas últimas duas camadas e tipo da camada top. A rede otimizada apresentou

melhores resultados que todas as demais avaliadas, obtendo 94% de acurácia e F1-score de 0,91. Os autores afirmam que, para a base em questão, CNNs menores apresentam um melhor resultado.

Barcelos (2017) conduziu um estudo utilizando a base de dados DRM-IR, focando em 70 pacientes. A abordagem concentrou-se na detecção de câncer de mama por meio de imagens de protocolo estático. O processo de classificação iniciou-se com a extração de características das pacientes, dividindo-as em três grupos: normal, com alterações benignas e com alterações malignas (câncer de mama).

O estudo gerou diversos subconjuntos com base nessas características, os quais foram submetidos a testes utilizando dois classificadores distintos: redes neurais artificiais e máquina de vetores de suporte (SVM). Em uma série de experimentos, o resultado mais promissor foi obtido com o uso de SVM como classificador, empregando uma função cúbica como função de kernel e uma abordagem *one vs all*. Os resultados revelaram uma acurácia de 80,95%, especificidade de 83,33% para pacientes normais, 85,71% para pacientes com alterações benignas e 75% de sensibilidade para pacientes com alterações malignas.

O estudo conduzido por Vogado et al. (2017) apresenta uma abordagem inovadora para o diagnóstico de Leucemia por meio da análise de imagens de sangue. A proposta utiliza três CNNs pré-treinadas, nomeadamente a AlexNet, a CaffeNet e a VggNet. O estudo não apenas realiza a avaliação individual do desempenho dessas redes, mas também explora a eficácia de um comitê de classificadores do tipo voto majoritário, que combina as saídas das CNNs para aumentar a confiabilidade dos resultados, categorizando as imagens como doentes ou não. A base de dados empregada para a condução dos experimentos é a ALL-IDB1, composta por um total de 108 imagens, das quais 59 são saudáveis e 49 são de pacientes doentes.

Os resultados obtidos nos experimentos demonstram um desempenho notável nas três redes testadas por Vogado et al. (2017). Em particular, nos testes com segmentação das imagens, as redes AlexNet e VggNet apresentaram resultados idênticos, alcançando uma acurácia impressionante de 96,29%. Em contrapartida, a rede CaffeNet obteve uma acurácia ligeiramente menor, registrando 94,44%. Contudo, destaca-se que o comitê de classificadores revelou um desempenho superior, atingindo uma acurácia de 97,22%. Esses resultados validam a eficácia do uso de comitês de classificadores para tornar a classificação mais confiável e robusta, proporcionando *insights* valiosos para aplicações clínicas futuras.

3 Desenvolvimento

Neste capítulo serão descritas algumas etapas realizadas para garantir o melhor desempenho durante os experimentos, baseados nos métodos anteriormente descritos. Na Figura 7 é possível visualizar um fluxograma que ilustra as etapas dos métodos aplicados.

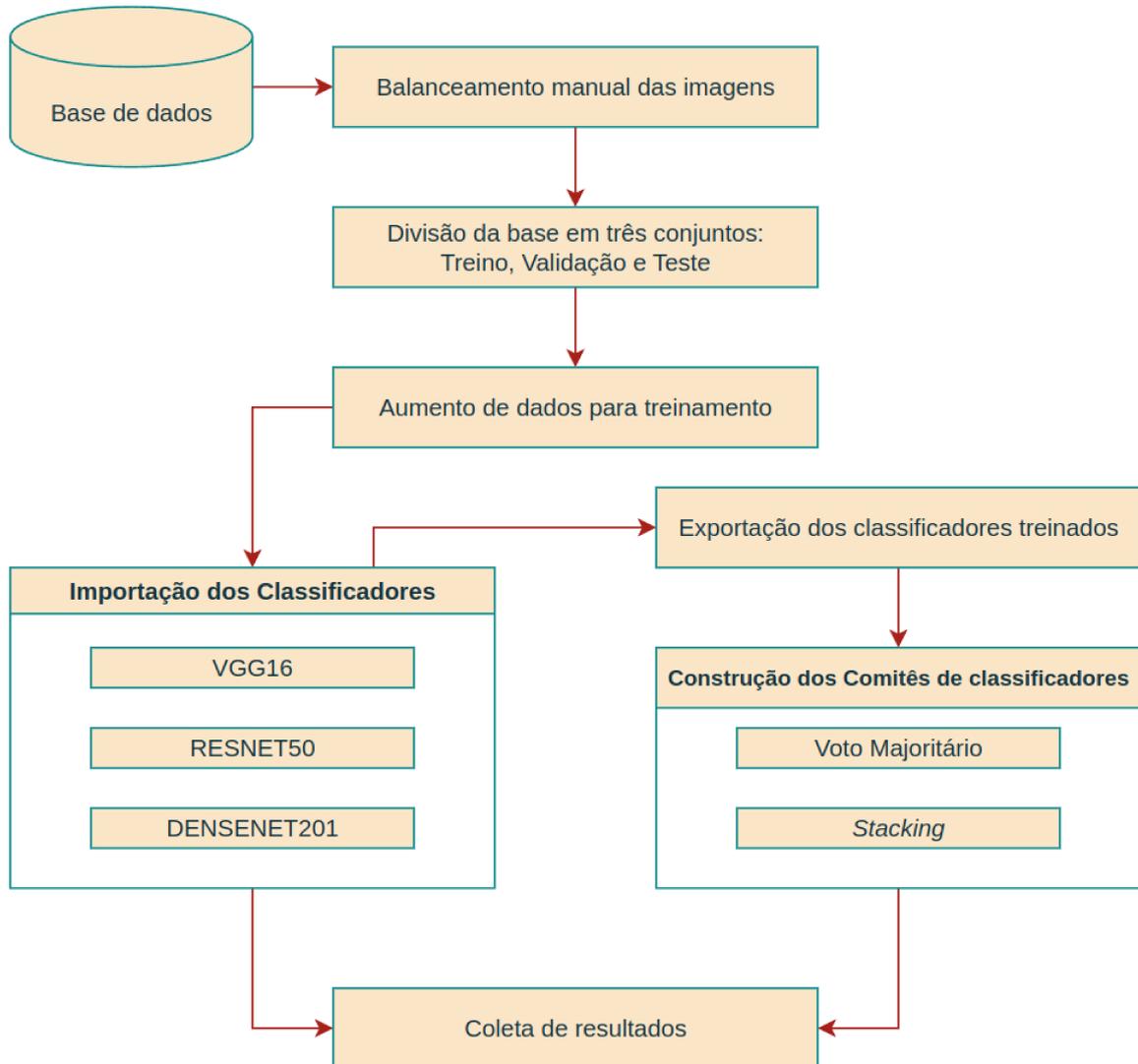


Figura 7 – Fluxograma das etapas abordadas

3.1 Ajuste na base de dados

Como descrito anteriormente, a base de dados foi sendo atualizada ao longo do tempo. No momento da elaboração deste trabalho, a base conta com 157 pacientes saudáveis e 38 doentes. Observa-se que a base está desbalanceada, e devido ao número limitado de amostras, optou-se pelo balanceamento manual das imagens. Essa fase envolveu a seleção criteriosa de pacientes saudáveis com menor presença de ruído nas imagens, enquanto todas as amostras de pacientes doentes foram preservadas, resultando assim em uma igualdade na quantidade de imagens de ambas as condições.

O particionamento do conjunto de dados em três subconjuntos - Treino, Validação e Teste - foi executado de maneira automatizada durante a implementação dos algoritmos utilizando a biblioteca *Scikit Learn*. O conjunto de treinamento engloba 70% das imagens, enquanto os conjuntos de validação e teste correspondem a 15% cada. Estes valores foram escolhidos dada a quantidade muito baixa de amostras, afim de se evitar o *overfitting* e para avaliar a capacidade de generalização dos modelos utilizados. A função responsável pela divisão dos subconjuntos requer um número aleatório, o qual foi fixado nas implementações para garantir consistência nas divisões, mantendo o mesmo padrão para todas os experimentos. Essa abordagem visa assegurar a reprodutibilidade e a comparabilidade dos resultados ao longo dos diferentes experimentos realizados.

3.2 Aumento de dados

O enriquecimento de dados foi incorporado dinamicamente durante o processo de treinamento dos classificadores, implicando que cada imagem, ao ser instanciada, possa ser submetida a uma ou mais transformações predefinidas. Essa prática de aumento de dados visa atenuar o risco de *overfitting*, especialmente considerando que a base de dados utilizada apresenta uma quantidade limitada de imagens.

Para ampliar a diversidade do conjunto de treinamento, optou-se por duas estratégias de aumento de dados: rotação aleatória da imagem, variando de -10° a 10° , e inversão horizontal, ambas aplicadas de forma aleatória. Essas técnicas contribuem para fortalecer a capacidade do modelo em generalizar padrões e comportamentos, promovendo assim um desempenho mais robusto diante de novos dados.

3.3 Classificação com as redes pré-treinadas

Cada arquitetura de rede neural - VGG16, RESNET50 e DENSENET201 - passou por um treinamento individualizado. Após extensivos testes, decidiu-se manter os parâmetros de treinamento consistentes em todos os experimentos. Os resultados mais promissores foram alcançados ao utilizar 30 épocas para o treinamento, adotando a função

de perda *Cross Entropy*, e empregando o otimizador *Adam* com uma taxa de aprendizado de 0,00001 e um *weight decay* de 0,0001.

Dada a restrição de uma quantidade limitada de imagens, a fase de validação demonstrou pequenas oscilações nos resultados, se comparada ao treinamento. Visando a construção de comitês de classificadores mais robustos, foi adotada a prática de salvar os modelos em treinamentos específicos, optando por aqueles em que a etapa de validação evidenciou as menores oscilações.

Nas Figuras 8 e 9 é possível observar a oscilação no processo de validação citada acima. Cada gráfico foi gerado após um treinamento distinto da rede VGG16. A palavra *Test* apresentada nas legendas foi colocada erroneamente e pode gerar confusão, já que se trata da validação, ou seja, deveria ter o nome de *Validation*.

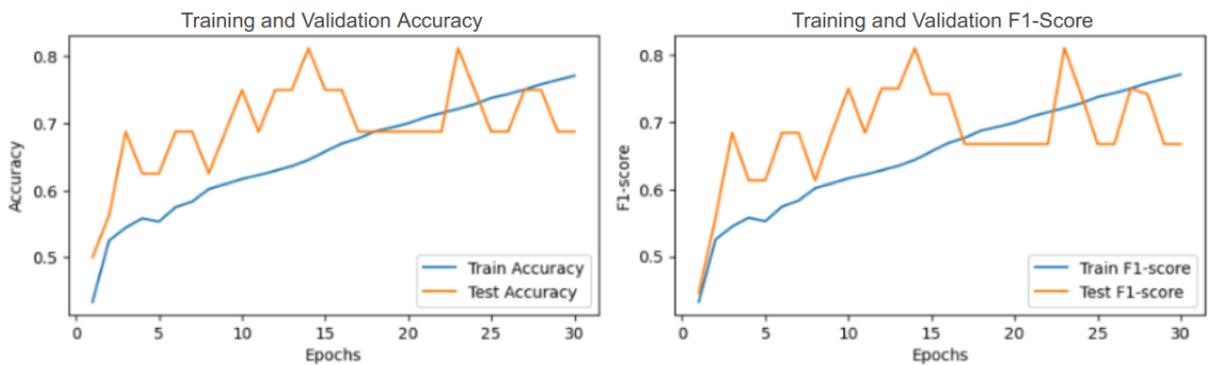


Figura 8 – Gráfico gerado após treinamento em um dos experimentos com a rede VGG16

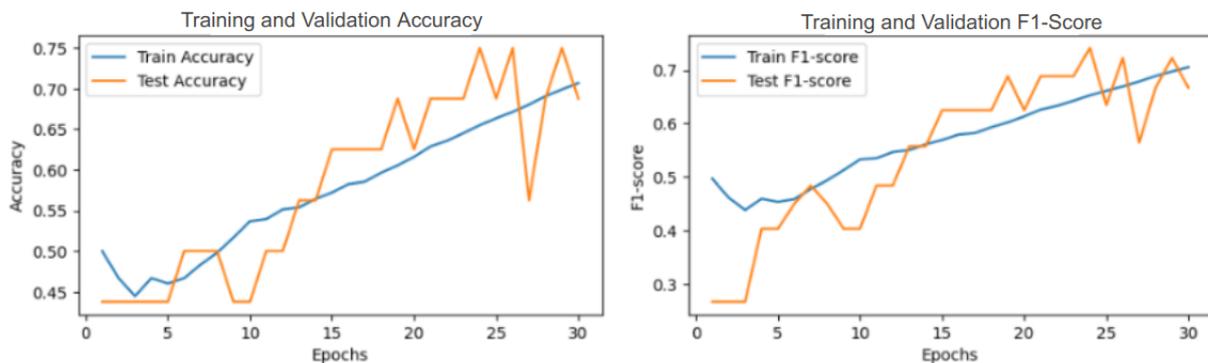


Figura 9 – Gráfico gerado após treinamento em um dos experimentos com a rede VGG16, a critério de comparação

4 Resultados e Discussão

Todos os experimentos foram desenvolvidos na plataforma de colaboração do Google (Colab) utilizando a linguagem Python. Colab é um serviço hospedado do *Jupyter Notebook* que não requer configuração para ser usado e fornece acesso gratuito a recursos de computação, incluindo GPUs e TPUs. O Colab é especialmente adequado para aprendizado de máquina e ciência de dados.

A avaliação do desempenho dos modelos foi conduzida mediante a aplicação das previsões sobre o conjunto de testes, o qual possui amostras diferentes dos conjuntos de teste e validação, ou seja, não possui ligação com o processo de treinamento. Os resultados apresentados são a média de 10 execuções distintas de cada modelo. Essa abordagem visa proporcionar uma visão robusta e estável da eficácia do modelo, incorporando a variabilidade inerente aos dados e ao processo de treinamento. Os resultados obtidos durante esta etapa do projeto podem ser visualizadas na Tabela 1.

	Acurácia(%)	Sensibilidade(%)	Especificidade(%)	F1-Score(%)
VGG16	70,92	83,33	56,00	75,67
RESNET50	66,37	60,00	74,00	63,10
DENSENET201	66,37	83,33	46,00	72,66

Tabela 1 – Média dos resultados obtidos em 10 execuções para cada modelo

Ao comparar os resultados da Tabela 1 com estudos correlatos, evidencia-se uma disparidade em favor de alguns trabalhos, os quais alcançaram desempenhos mais satisfatórios. Contudo, é crucial ressaltar que a proposta central desta pesquisa reside na comparação entre os resultados individuais de redes neurais e os resultados de um comitê que integra as predições de cada uma dessas redes para a tarefa de classificação. Nesse contexto, optou-se por não empregar técnicas adicionais que poderiam aprimorar significativamente os resultados, tais como técnicas de melhoramento de imagem e segmentação da região de interesse. Mesmo sem a implementação dessas técnicas, é notável que os resultados alcançados nas redes neurais individuais utilizadas neste trabalho se aproximam de alguns resultados obtidos por outros pesquisadores, indicando a viabilidade e eficácia da abordagem proposta.

	Acurácia(%)	Sensibilidade(%)	Especificidade(%)	F1-Score(%)
Voting	90,91	100,00	80,00	92,31
Stacking	54,55	100,00	00,00	70,59

Tabela 2 – Resultados obtidos por cada comitê de classificadores

Na implementação do primeiro comitê de classificadores - Voto Majoritário - inicialmente procedeu-se com o carregamento de cada modelo de rede neural previamente treinado. Subsequentemente, foram coletadas as predições de cada um desses modelos sobre o conjunto de testes. As predições de todos os modelos foram consolidadas em uma matriz, onde cada coluna representa as previsões de um modelo específico. Em seguida, para cada linha dessa matriz, realizou-se a análise da predição mais comum, ou seja, a votação majoritária, cujo resultado foi registrado em uma lista.

Esta lista contém as predições finais para cada imagem avaliada. A avaliação do desempenho do comitê é conduzida ao comparar essas predições finais com os rótulos verdadeiros associados a cada imagem no conjunto de teste. Em comparação com os resultados obtidos pelas redes neurais individualmente, observa-se uma melhora significativa ao empregar o comitê de classificadores do tipo voto majoritário. Os resultados obtidos são apresentados na Tabela 2, e mostram um impressionante potencial nesta abordagem, sendo uma acurácia de 90,91%, sensibilidade de 100%, especificidade de 80% e F1-Score de 92,31%. Esses resultados consolidados sugerem que a abordagem do comitê, ao integrar as previsões das redes neurais, resultou em melhorias substanciais, proporcionando uma avaliação mais robusta e abrangente do sistema.

No contexto do comitê do tipo *Stacking*, a coleta de predições foi realizada sobre o conjunto de treino, escolha motivada pela escassez de imagens, tornando este o conjunto mais abundante em amostras. As previsões obtidas para cada imagem são verticalmente empilhadas, permitindo uma análise comparativa das saídas de cada modelo. Esse procedimento resulta na criação de uma espécie de base de dados, a qual servirá como entrada para um meta-classificador. Posteriormente, o meta-classificador é treinado de maneira convencional, gerando novas previsões.

Para avaliar o desempenho desse comitê, as previsões das redes individuais são novamente coletadas, desta vez no conjunto de teste, sendo seguido por um novo empilhamento dessas previsões. Este conjunto empilhado é utilizado para avaliar o modelo. Com os resultados desta abordagem de comitê do tipo *Stacking*, apresentados na Tabela 2 é possível avaliar seu desempenho, destacando que este comitê não apresentou resultados satisfatórios, pois seu desempenho foi inferior ao obtido pelas redes neurais separadamente, tendo como resultados de acurácia, sensibilidade, especificidade e F1-Score, respectivamente, 54,55%, 100%, 0% e 70,59% .

Analisando os resultados acima é possível perceber que o comitê do tipo *Stacking* foi capaz de identificar corretamente todos os casos positivos (VP) em relação ao total de casos positivos, ou seja, o modelo não deixa de detectar nenhuma caso verdadeiramente positivo. Por outro lado, a especificidade extremamente baixa indica que o comitê não consegue distinguir corretamente os casos negativos (VN) do total de casos negativos. Em outras palavras, o comitê classificou erroneamente todos os casos negativos como positivos

ou não consegue identificar corretamente nenhum caso verdadeiramente negativo.

A utilização das previsões individuais de cada rede neural como entrada para o meta-classificador pode estar apresentando desafios significativos na captura das nuances do conjunto de dados. É possível que o meta-classificador esteja sendo influenciado de maneira desproporcional pelas previsões de uma ou mais redes, resultando em um impacto desigual na tomada de decisões finais. Essa influência desproporcional pode ser resultado da base de dados original que é extremamente pequena, o que resulta em um número limitado de entradas para o meta-classificador.

A escassez de dados impõe uma complexidade adicional ao aprendizado do modelo, que precisa lidar com uma pilha de previsões binárias (zeros e uns). Esse contexto de aprendizado em um espaço tão restrito pode dificultar a capacidade do meta-classificador em generalizar eficientemente novas instâncias, contribuindo assim para os resultados baixos observados. Portanto, considerar estratégias para aumentar o desempenho dos classificadores individualmente, resultando em saídas mais padronizadas, além de explorar alternativas no design do meta-classificador, pode ser essencial para melhorar o desempenho do comitê do tipo stacking.

5 Conclusão

O uso de imagens termográficas para a detecção do câncer de mama emerge como uma técnica promissora no auxílio ao diagnóstico dessa doença. Esta abordagem visa proporcionar uma alternativa menos invasiva e indolor para as pacientes, além de apresentar benefícios como custo reduzido e a capacidade de detectar a doença em mamas mais densas. Essa técnica vem sendo estudada a algum tempo, tendo trabalhos que apresentaram ótimos resultados em seus testes.

Ao comparar os resultados obtidos por cada CNN individualmente com estudos correlatos, nota-se uma certa disparidade. No entanto, é crucial ressaltar que a principal proposta deste trabalho foi a comparação entre os resultados individuais de redes neurais e os resultados de um comitê que integra as previsões de cada uma dessas redes.

No contexto do comitê de voto majoritário, destaca-se uma melhora significativa em relação às redes neurais individuais, evidenciada por uma acurácia de 90,91%, sensibilidade de 100%, especificidade de 80%, e F1-Score de 92,31%. Esses resultados enfatizam a eficácia do comitê ao integrar as previsões das redes neurais, proporcionando uma avaliação mais robusta e abrangente do sistema.

Por outro lado, ao explorar o comitê do tipo *stacking*, que utiliza um meta-classificador, observou-se um desempenho inferior em comparação com as redes neurais individuais. Uma análise detalhada sugere possíveis razões para esse desempenho, abrindo espaço para ajustes e otimizações que poderiam potencialmente maximizar a eficácia desse tipo de comitê.

Em suma, os resultados obtidos, apesar de apresentarem desafios em determinados cenários, sugerem que a abordagem de comitês pode ser uma estratégia promissora para aprimorar o desempenho na tarefa de classificação de imagens para detecção do câncer de mama, além de abrir espaço para que vários outros trabalhos possam se basear nesse estudo e fazer testes com os diversos tipos de comitês de classificadores existentes. No entanto, ressalta-se a necessidade de uma cuidadosa consideração dos aspectos específicos do conjunto de dados e do processo de treinamento para otimizar a eficácia dessa abordagem.

Este trabalho foi enriquecido pela aplicação de conhecimentos adquiridos em disciplinas fundamentais ao longo da graduação. Disciplinas como programação procedimental e estrutura de dados forneceram a base técnica necessária para o desenvolvimento eficiente deste projeto. Além disso, os conceitos aprendidos em inteligência computacional desempenharam um papel crucial na abordagem e implementação de algoritmos relevantes.

A disciplina de multimídias também contribuiu significativamente, proporcionando *insights* valiosos sobre técnicas de melhoria de imagens que poderiam ter sido aplicadas neste trabalho, ou que podem ser aplicadas em trabalhos futuros. Essa interdisciplinaridade demonstra a relevância de uma formação sólida e abrangente, destacando a importância não apenas da especialização em uma área específica, mas também da integração de conhecimentos de diversas disciplinas da graduação.

Referências

- ANDRADE, F.; PAIVA, A.; CORREA, A. Análise de imagens de termografia dinâmica para classificação de alterações na mama usando séries temporais. In: **Conference on Graphics, Patterns and Images is an international conference annually promoted by the Brazilian Computer Society**. Niterói, RJ, Brazil: [s.n.], 2017. v. 30, p. 1–4. Disponível em: <<http://sibgrapi2017.ic.uff.br/>>. Citado na página 21.
- AZEVEDO, W. et al. Seleção de atributos para apoio ao diagnóstico do câncer de mama usando imagens termográficas, algoritmos genéticos e otimização por enxame de partículas. In: **Anais do II SIMPÓSIO DE INOVAÇÃO EM ENGENHARIA BIOMÉDICA**. Recife, Brazil: Sociedade Brasileira de Engenharia Biomédica, 2018. Citado na página 10.
- BARATLOO, A.; HOSSEINI, M.; NEGIDA, A.; ASHAL, G. E. Part 1: simple definition and calculation of accuracy, sensitivity and specificity. **ARCHIVES OF ACADEMIC EMERGENCY MEDICINE (EMERGENCY)**, 2015. Citado na página 14.
- BARCELOS, C. G. Detecção de câncer de mama utilizando imagens termográficas. **TCC, Universidade Federal de Uberlândia, Brasil**, 2017. Disponível em: <<https://repositorio.ufu.br/handle/123456789/20255>>. Citado 6 vezes nas páginas 10, 11, 16, 17, 18 e 23.
- BORCHARTT, T. B. Análise de imagens termográficas para a classificação de alterações na mama. **PhD diss., Fluminense Federal University, Brazil**, 2013. Citado 3 vezes nas páginas 9, 10 e 18.
- CARDOSO, J. P. S. Classificação de pacientes em saudáveis e não saudáveis em termografia infravermelha dinâmica de mamas usando séries temporais. **Universidade Federal do Maranhão, Brasil**, 2018. Disponível em: <<http://hdl.handle.net/123456789/3550>>. Citado 3 vezes nas páginas 10, 11 e 22.
- CHAVES, E. d. L. Detecção de câncer de mama por meio de imagens infravermelhas utilizando redes neurais convolucionais. **Universidade Federal de Uberlândia**, 2019. Disponível em: <<https://repositorio.ufu.br/handle/123456789/25938>>. Citado 5 vezes nas páginas 5, 11, 18, 19 e 22.
- FORCELINI, F. et al. Termografia infravermelha aplicada ao design: protocolo de coleta de dados termográficos para o desenvolvimento de projetos. 2019. Disponível em: <<https://repositorio.ufsc.br/handle/123456789/214903>>. Citado na página 17.
- HAYKIN, S. **Neural networks: a comprehensive foundation**. Upper Saddle River, New Jersey: Prentice Hall PTR, 1998. Citado na página 20.
- IBARRA-CASTANEDO, C. Quantitative subsurface defect evaluation by pulsed phase thermography: depth retrieval with the phase. 2005. Citado 2 vezes nas páginas 5 e 18.
- INCA. **I.N. de C. Tipos de Câncer: Câncer de Mama**. 2023. Disponível em: <<https://www.gov.br/inca/pt-br/assuntos/cancer/tipos/mama>>. Accessed: 2023-04-07. Citado 4 vezes nas páginas 5, 9, 16 e 17.

SILVA, L.; SAADE, D.; SEQUEIROS, G.; SILVA, A.; PAIVA, A.; BRAVO, R.; CONCI, A. A new database for breast research with infrared image. **Journal of Medical Imaging and Health Informatics**, American Scientific Publishers, v. 4, n. 1, p. 92–100, 2014. Disponível em: <<https://doi.org/10.1166/jmihi.2014.1226>>. Citado na página 12.

SILVA, L. A. d. **Classificação de nódulos mamográficos utilizando um comitê de redes neurais artificiais**. Tese (Doutorado) — Universidade de São Paulo, 2005. Disponível em: <<https://doi.org/10.11606/D.3.2005.tde-05092006-203231>>. Citado na página 20.

SILVA, W. K. N. d. Construções de comitês de classificadores multirrótulos no aprendizado semissupervisionado multidescrição. Universidade Federal Rural do Semi-Árido, 2017. Disponível em: <<https://repositorio.ufersa.edu.br/handle/tede/812>>. Citado 3 vezes nas páginas 11, 20 e 21.

SOCIETY, A. C. **Breast Density and Your Mammogram Report**. 2023. Disponível em: <<https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/mammograms/breast-density-and-your-mammogram-report.html>>. Accessed: 2023-06-27. Citado 3 vezes nas páginas 5, 9 e 10.

VOGADO, L. H.; VERAS, R. M.; ANDRADE, A. R.; SANTOS, L. G.; AIRES, K. R.; MACHADO, V. P. Um sistema de diagnóstico de leucemia utilizando cnn's pré-treinadas e um comitê de classificadores. In: SBC. **Anais do XVII Workshop de Informática Médica**. 2017. Disponível em: <<https://doi.org/10.5753/sbcas.2017.3719>>. Citado 2 vezes nas páginas 19 e 23.

ZULUAGA-GOMEZ, J.; MASRY, Z. A.; BENAGGOUNE, K.; MERAGHNI, S.; ZERHOUNI, N. A cnn-based methodology for breast cancer diagnosis using thermal images. **Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization**, Taylor & Francis, v. 9, n. 2, p. 131–145, 2021. Disponível em: <<https://doi.org/10.1080/21681163.2020.1824685>>. Citado 2 vezes nas páginas 11 e 22.