

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Thiago Mota Carvalho

**Explorando Modelos em Cascata de CNNs e
Fractais para Classificação de Displasia Oral**

Uberlândia, Brasil

2023

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Thiago Mota Carvalho

**Explorando Modelos em Cascata de CNNs e Fractais
para Classificação de Displasia Oral**

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

Orientador: Marcelo Zanchetta do Nascimento

Universidade Federal de Uberlândia – UFU

Faculdade de Computação

Bacharelado em Ciência da Computação

Uberlândia, Brasil

2023

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

C331 2023	<p>Carvalho, Thiago Mota, 1998- Explorando Modelos em Cascata de CNNs e Fractais para Classificação de Displasia Oral [recurso eletrônico] / Thiago Mota Carvalho. - 2023.</p> <p>Orientador: Marcelo Zanchetta do Nascimento. Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Uberlândia, Graduação em Ciência da Computação. Modo de acesso: Internet. Inclui bibliografia. Inclui ilustrações.</p> <p>1. Computação. I. Nascimento, Marcelo Zanchetta do, 1976-, (Orient.). II. Universidade Federal de Uberlândia. Graduação em Ciência da Computação. III. Título.</p> <p style="text-align: right;">CDU: 681.3</p>
--------------	---

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Faculdade de Computação

Av. João Naves de Ávila, nº 2121, Bloco 1A - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902

Telefone: (34) 3239-4144 - <http://www.portal.facom.ufu.br/> facom@ufu.br



ATA DE DEFESA - GRADUAÇÃO

Curso de Graduação em:	Bacharelado em Ciência de Computação				
Defesa de:	Projeto de Graduação 2				
Data:	30/11/2023	Hora de início:	20:15	Hora de encerramento:	21:15
Matrícula do Discente:	11811BCC026				
Nome do Discente:	Thiago Mota Carvalho				
Título do Trabalho:	Explorando Modelos em Cascata de CNNs e Fractais para Classificação de Displasia Oral				
A carga horária curricular foi cumprida integralmente?	<input checked="" type="checkbox"/> Sim <input type="checkbox"/> Não				

Reuniu-se, remotamente, a Banca Examinadora, designada pelo Colegiado do Curso de Graduação em Ciência da Computação, assim composta: Professores (as): João Henrique de Souza Pereira - FACOM/UFU; Luiz Cláudio Theodoro - FACOM/UFU; e Marcelo Zanchetta do Nascimento - FACOM/UFU, orientador do candidato.

Iniciando os trabalhos, o presidente da mesa, Dr. Marcelo, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao discente a palavra, para a exposição do seu trabalho. A duração da apresentação do discente e o tempo de arguição e resposta foram conforme as normas do curso.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

(X) Aprovado(a) Nota [95].

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Marcelo Zanchetta do Nascimento, Professor(a) do Magistério Superior**, em 30/11/2023, às 21:14, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Luiz Cláudio Theodoro, Professor(a) do Magistério Superior**, em 30/11/2023, às 21:15, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **João Henrique de Souza Pereira, Professor(a) do Magistério Superior**, em 30/11/2023, às 21:18, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4997854** e o código CRC **DD362123**.

Referência: Processo nº 23117.083629/2023-70

SEI nº 4997854

Agradecimentos

Agradeço, em especial, ao professor Marcelo, pela paciência na orientação, pelo tempo dedicado e pela compreensão ao longo de todo o projeto. Agradeço, também, em especial, ao professor João Henrique pelo empenho em criar uma cultura de maratona de programação na UFU e por ter me introduzido no ecossistema de competições. Agradeço, também, ao Igor e ao Mateus, minha graduação não seria a mesma sem vocês. Registro meus agradecimentos, ainda, a todos os amigos que estiveram presente, de alguma forma, compartilhando experiências e dando forças a cada dia desde o início da minha graduação.

Resumo

A classificação de imagens histológicas é uma tarefa que tem sido amplamente explorada nas recentes pesquisas de visão computacional. A abordagem mais estudada para esta tarefa tem sido a aplicação de aprendizado profundo por meio de modelos de Aprendizagem Profunda (AP). Entretanto, o uso de AP no contexto da classificação de imagens histológicas tem ainda algumas limitações, como a necessidade de grandes conjuntos de dados e a dificuldade de implementar um modelo generalizado capaz de classificar diferentes tipos de tecido histológico. Este trabalho, se propõe a investigar classificadores e comitês baseados em representações de atributos fractais em modelos de rede AP em cascata para avaliação de lesões da cavidade oral. A extração das características foi obtida com atributos fractais do conjunto de dados histológicos. Esses atributos são reorganizados em uma matriz a fim compor uma imagem de representação. Foram avaliados dois métodos, o primeiro consistiu em utilizar as imagens histológicas originais como entrada para os classificadores. O segundo método consistiu em fazer uso das representações fractais como entrada para os classificadores. Fazendo uso de representações fractais, foram obtidos valores da métrica acurácia que variam de 98,75% a 100,00%. Quando aplicado apenas as imagens histológicas de displasia, foram obtidas acurácias que variam de 85,89% a 96,61%. Os resultados foram relevantes e mostram que essa metodologia pode ser utilizada como forma de auxiliar especialistas da área da saúde para diagnóstico de displasia.

Palavras-chave: Displasia, CAD, *ResNet*, *MobileNet*, Comitê, Fractal, CNN em Cascata

Lista de ilustrações

Figura 1 – Exemplos de carcinoma celular escamosa na língua (a) e com invasão na gengiva (b). Fonte: (KADEMANI, 2019).	16
Figura 2 – Exemplos de tecido com displasia: a) representa um tecido com displasia leve, b) um exemplo de displasia moderada e c) amostra de displasia severa.	17
Figura 3 – Etapas empregadas em um sistema CAD para auxiliar um especialista.	17
Figura 4 – Imagem de uma árvore sob duas escalas de observação diferentes: imagem completa (a) e imagem observada sob uma escala 3 vezes menor (b).	20
Figura 5 – Simulação do algoritmo <i>gliding-box</i> em uma imagem com dimensões 8×8 . Fonte: (ROBERTO et al., 2021)	21
Figura 6 – Atribuindo os valores aos pixels em uma caixa 3×3 depois de calcular o valor da distância Δ . Fonte: (ROBERTO et al., 2021)	22
Figura 7 – Tapetes de <i>Sierpinski</i> . As imagens (a) e (b) possuem similares valores FD, contudo eles possuem valores diferentes de lacunaridade. Fonte: (MANDELROT, 2015)	23
Figura 8 – Uma matriz com dimensões de 15×15 , em (a) a percolação se manifesta devido à presença de um aglomerado que se estende desde uma extremidade até a outra da matriz, sendo visualizado em uma tonalidade de verde claro (b). Fonte: (ROBERTO et al., 2021).	24
Figura 9 – Exemplo de como ocorre uma convolução.	26
Figura 10 – Exemplo de abordagens de <i>max pooling</i> e <i>average pooling</i>	27
Figura 11 – Demonstração de uma transformação feita na camada <i>flatten</i>	28
Figura 12 – Estrutura da arquitetura <i>AlexNet</i> Fonte: (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).	28
Figura 13 – Exemplo de um Árvore de decisão.	29
Figura 14 – Exemplo de uma floresta aleatória.	30
Figura 15 – Exemplo de aumento de dados utilizando a técnica de rotação em 90 graus: imagem original (a) e imagem rotacionada (b).	32
Figura 16 – Ilustração de uma etapa de um modelo usando camada <i>dropout</i> com $p = 0,5$. Os vértices em cor cinza foram desativados na iteração e os outros continuaram ativos.	34
Figura 17 – Exemplo de um Comitê com N classificadores.	34
Figura 18 – Sequência metodológica utilizada para a classificação de lesões da cavidade oral.	39

Figura 19 – Exemplos de tecidos histológicos da cavidade bucal digitalizados: (a) saudável, (b) displasia leve, (c) displasia moderada e (d) displasia severa. Fonte : (SILVA et al., 2019)	41
Figura 20 – Método de gráfico de recorrência sendo aplicado ao vetor de características para gerar as imagens. Fonte: (ROBERTO et al., 2021)	43
Figura 21 – Estrutura da arquitetura da <i>Resnet50</i> pré-treinada com <i>ImageNet</i> Fonte : (TANG; LI; WANG, 2020)	44
Figura 22 – Estrutura da arquitetura da <i>MobileNet</i> Fonte : (PHIPHITPHATPHAISIT; SURINTA, 2020)	44
Figura 23 – O comitê formado por 3 algoritmos (Algoritmo 1, 2, 3), o qual recebe o resultado de cada modelo e realiza a operação estatística (moda ou média).	48
Figura 24 – Treinamento em 100 épocas utilizando a rede <i>MobileNet</i> com as imagens obtidas pelo método GR.	50
Figura 25 – Gráficos que mostram o processo de treinamento da (a) <i>ResNet50</i> e (b) <i>ResNet50V2</i> utilizando as imagens obtidas pelo método ID.	51

Lista de tabelas

Tabela 1 – Estrutura do modelo proposto customizado com as camadas e saída das operações em cada camada.	45
Tabela 2 – Conjunto de dados investigados nesse trabalho para classificação das lesões.	46
Tabela 3 – Número de imagens após aplicação das operações de aumento de dados.	46
Tabela 4 – Desempenho dos Algoritmos de ML para Classificação das Lesões Histológicas	49
Tabela 5 – Avaliação das imagens com uso dos modelos de CNNs em cascata. . . .	50
Tabela 6 – Desempenho dos modelos baseados em comitê de classificadores para as abordagens ML e CNN.	52
Tabela 7 – Avaliação da metodologia proposta em relação às abordagens existentes para classificação de imagens orais presentes na literatura.	52

Lista de abreviaturas e siglas

ac	Acurácia
AUC	Area under ROC Curve
AM	Aprendizado de Máquina
AP	Aprendizado Profundo
CAD	Computer-Aided Diagnosis
CNN	Convolutional Neural Network
DT	Decision Tree
f1	F1-Score
FD	Fractal Dimension
FE	Feature Engineering
FN	Falso Negativo
FP	Falso Positivo
GR	Gráfico de Recorrência
GLCM	Grey level co-occurrence matrix
H&E	Hematoxilina-Eosina
HOP	Hermite Orthogonal Polynomial
ID	Imagens Digitais (ID) histológicas de displasia
LAC	Lacunarity
LF	Learning Features
LR	Logistic Regression
MN	MobileNet
MF	Morphological Features
ORB	Oriented FAST and Rotated BRIEF

OMS	Organização Mundial de Saúde
PERC	Percolation
RF	Random Forest
RGB	Sistema de cores Red, Green e Blue
ROI	Region of Interest
RN50	ResNet50
RN50V2	ResNet50V2
RP	Recurrence Plot
sens	Sensibilidade
SVM	Support Vector Machine
TA	Texture Analysis
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

Sumário

1	INTRODUÇÃO	12
1.1	Justificativa	13
1.2	Objetivos	13
1.2.1	Objetivo Geral	13
1.2.2	Objetivos específicos	14
1.3	Organização deste Trabalho de Conclusão de Curso	14
2	FUNDAMENTAÇÃO TEÓRICA	15
2.1	Câncer	15
2.2	Displasia	16
2.3	Sistema de Diagnóstico Auxiliado por Computador	17
2.4	Análise de Textura	18
2.5	Geometria Fractal	19
2.5.1	Dimensão Fractal	22
2.5.2	Lacunaridade	23
2.5.3	Percolação	24
2.6	Classificação	25
2.6.1	Redes Neurais Convolucionais	26
2.6.2	Árvore de Decisão	29
2.6.3	Floresta Aleatória	29
2.6.4	Regressão Logística	31
2.7	Aumento de dados	31
2.8	Regularizadores	32
2.9	Comitê de Classificadores	34
2.10	Métricas de avaliação	36
2.11	Trabalhos relacionados	37
3	DESENVOLVIMENTO	39
3.1	Visão geral	39
3.2	Banco de Imagens	40
3.3	Representação 2D de Atributos Fractais	41
3.3.1	Geometria Fractal	41
3.3.2	Representação com Método de Gráfico de Recorrência	42
3.4	Rede Neural Convolucional	42
3.5	Operações de Aumento de Dados	46
3.6	Comitê de Classificadores	47

3.7	Avaliação de Desempenho das Abordagens	48
4	RESULTADOS	49
4.1	Avaliação de Desempenho dos Algoritmos de ML para Lesões Histológicas	49
4.2	Resultados com os Modelos CNNs em Cascata	50
4.3	Investigação dos Comitês de Classificadores	51
4.4	Análise de Desempenho com Estudos da Literatura	52
5	CONCLUSÃO	53
	REFERÊNCIAS	54

1 Introdução

A displasia é considerada uma das principais lesões pré-cancerosas. As pré-cânceres são condições pré-malignas, lesões ou tumores que provocam a multiplicação celular, elevando o risco da manifestação de algum câncer (WARNAKULASURIYA et al., 2008). Por esse motivo, o diagnóstico deve ser feito precocemente para possibilitar melhores chances de sucesso no tratamento do paciente. Uma das formas de fazer o diagnóstico é utilizando sistemas computacionais de apoio ao diagnóstico (do inglês, *computer aided diagnosis - CAD*), que com uso de inteligência artificial, é possível fazer a identificação das lesões em amostras.

Segundo a Organização Mundial de Saúde (OMS), o câncer oral é classificado como o 13º câncer mais comum em todo o mundo. No ano de 2020, a incidência global foi de cerca de 377 mil novos casos e mais de 170 mil mortes. O câncer oral inclui os cânceres de lábio, outras partes da boca e orofaringe. Os números da ocorrência de câncer de lábio e cavidade oral foi estimada em 377.713 novos casos e 177.757 mortes em 2020. O câncer bucal é mais comum em homens e idosos, mais mortal em homens em comparação com mulheres e varia fortemente de acordo com circunstâncias socioeconômicas. Além disso, as displasias apresentam entre 6% e 36% de chance de evoluírem o estado para câncer (SMITH et al., 2009).

O uso de sistemas CAD pode auxiliar o especialista e contribuir para uma segunda opinião para a tomada de decisão durante o diagnóstico. Esses sistemas são compostos por métodos de processamento de imagens, reconhecimento de padrões e inteligência artificial. Na primeira etapa, chamada de pré-processamento, é realizada a correção de distorções geométricas e eliminação de ruídos. Após isso, um realce da imagem pode ser empregado para permitir que os objetos sejam melhor discriminados entre si. A etapa de segmentação busca subdividir a imagem em objetos, possibilitando que ela seja descrita por meio dos objetos resultantes. Após a etapa da segmentação, é necessário o uso de algoritmos para extração de características para computar, de maneira eficaz, valores que descrevem os dados presentes numa imagem. Dentre os métodos de extração presentes na literatura, há duas formas, são elas, os métodos de extração de características aprendidas (do inglês, *learning features - LF*) e os métodos que exploram a extração de características por métodos tradicionais (do inglês, *feature engineering - FE*). Por fim, nesse sistema se tem a etapa de classificação em que as características são rotuladas em classes.

Os métodos FE podem usar informações internas dos objetos como textura, intensidade de brilho, etc. Um dos métodos FE é a matriz de coocorrência (do inglês, *gray level co-occurrence matrix - GLCM*) que consiste numa representação de probabilidades de

ocorrência de intensidade de brilho na imagem (MOHANAI AH; SATHYANARAYANA; GURUKUMAR, 2013). Nos métodos LF, as características são automaticamente extraídas por redes neurais convolucionais (do inglês, *convolutional neural network* - CNN). Existem diversos modelos de CNNs, das quais tem-se a *AlexNet*, a primeira rede neural a usar unidades processamento gráfico (do inglês, *graphics processing unit* - GPU) para melhorar o desempenho no processamento. Outros modelos, tais como, *MobileNet*, *ResNet* e *DenseNet* estão disponíveis atualmente na literatura. Com crescente capacidade de processamento dos computadores, a aplicação da CNN permitiu avanços significativos (LECUN; BENGIO; HINTON, 2015). De maneira complementar, o treinamento de redes neurais convolucionais tornou-se viável (LE et al., 2012), pois a principal estratégia apresenta a redução das taxas de erros de classificação, especialmente em problemas complexos. Isso ocorre devido ao aumento do número de camadas (profundidade) que define a arquitetura de uma rede.

O uso de estratégias que exploram modelos de CNN e representações fractais tornam-se relevantes em diversos domínios, mas ainda há desafios em aberto para classificação de amostras de tecidos da cavidade oral para os diversos graus de displasia.

1.1 Justificativa

A displasia é caracterizada pelo desenvolvimento celular fora do normal que pode resultar em formação de tecidos e órgãos em qualquer parte do corpo. Ademais, a displasia está muito ligada ao câncer, sendo considerada uma das principais lesões pré-cancerosas. As anormalidades da displasia possuem diferentes graus e características e, por esse motivo, se faz necessária a investigação e reconhecimento de cada um deles para um melhor tratamento dessas lesões, o que possibilita que graus com potenciais malignos sejam evitados.

Para fazer esse reconhecimento, a inteligência artificial e a visão computacional podem ser de grande ajuda, uma vez que o diagnóstico dessas patologias feitas por um profissional são trabalhosas, demandam tempo e estão sujeitas a erros. Com isso, um sistema CAD que faz o reconhecimento dessas patologias pode ajudar os especialistas no diagnóstico, além de proporcionar maior confiança ao profissional e ao paciente durante o tratamento.

1.2 Objetivos

1.2.1 Objetivo Geral

Essa pesquisa tem como objetivo investigar uma nova abordagem de classificação de displasia da cavidade oral com técnicas computacionais baseadas em modelos de redes

neurais convolucionais em cascata e representações de descritores fractais.

1.2.2 Objetivos específicos

Como objetivos específicos, a pesquisa tem as seguintes metas:

- Usar métodos de aprendizagem profunda em cascata e geometria fractal para extração de descritores;
- Analisar a classificação de imagens histológicas de displasia oral em diversos graus de evolução da lesão;
- Averiguar o uso de representações 2D de descritores fractais para tarefas de classificação;
- Explorar modelos de comitês para tarefas de classificação das lesões histológicas;

1.3 Organização deste Trabalho de Conclusão de Curso

Neste capítulo, foram apresentadas as considerações iniciais, motivação e objetivos do desenvolvimento deste trabalho. Os próximos capítulos possuem a seguinte organização:

- Capítulo 2: Introdz os principais conceitos para a compreensão e validação do estudo proposto e apresenta os trabalhos correlatos com a pesquisa;
- Capítulo 3: Delimita a proposta do trabalho e suas técnicas e metodologias desenvolvidas;
- Capítulo 4: Apresenta os resultados alcançados e as respectivas análises do método de classificação, além de um comparativo com a literatura;
- Capítulo 5: São apresentadas as contribuições obtidas, considerações finais e a motivação para trabalhos futuros.

2 Fundamentação Teórica

Este capítulo apresenta a descrição de conceitos que são fundamentais para o entendimento do domínio da pesquisa, assim como as principais técnicas exploradas. Além disso, são tratados os trabalhos relacionados do domínio de investigação, os quais exploraram ferramentas para auxiliar os especialistas no aprimoramento de diagnóstico das lesões investigadas.

2.1 Câncer

Câncer é o nome atribuído a um conjunto de doenças de desordem genética, as quais têm em comum o crescimento desordenado de células e são causadas por mutações no DNA. Essas doenças podem ser contraídas de forma espontânea ou devido a fatores ambientais. Um dos tipos comuns é o câncer da cavidade oral, quando o diagnóstico ocorre em estágio avançado apresenta 50% de chance de sobrevivência do indivíduo, considerando um diagnóstico tardio, superior a cinco anos (RAMOS, 2020). Nesse tipo, as células com mutações estão sujeitas a seleção natural, ou seja, há vantagens de crescimento e sobrevivência, podendo ter um crescimento desordenado em que parte dessas células dominam toda a estrutura comprometida (KUMAR et al., 2008). Quando há o diagnóstico das lesões do tipo carcinomas, a intervenção cirúrgica procura retirar regiões próximas à lesão, para garantir que a área comprometida não irá se manifestar em outras partes do corpo humano (KUMAR et al., 2008).

Além dos carcinomas, na cavidade oral podem ocorrer lesões pré-cancerosas, como a eritroplasia, a leucoplasia e a fibrose submucosa oral (TILAKARATNE et al., 2019). Essas lesões podem se apresentar junto às displasias epiteliais, que têm potencial a evoluir para o grau de severidade e ocorrer a transformação em carcinomas (KADEMANI, 2019). Por esse motivo, é de extrema importância a identificação precoce e a classificação da severidade das lesões, desse modo, permitindo um melhor tratamento ao paciente. Na Figura 1, tem-se exemplos de casos com a presença de lesão celular escamosa na região da língua e gengiva.



Figura 1 – Exemplos de carcinoma celular escamosa na língua (a) e com invasão na gengiva (b). Fonte: (KADEMANI, 2019).

2.2 Displasia

A displasia, um dos tipos de lesões pré-cancerosas, se caracteriza pela alteração morfológica da estrutura celular do tecido. Essa lesão, quando não tratada em estágio inicial, há chances de 6% a 36% de evoluir para um câncer (MEHANNA et al., 2009) (SMITH et al., 2009). Os pesquisadores, no estudo em Warnakulasuriya et al. (2008), mencionam os critérios que são utilizados para fazer classificação da displasia:

- Estratificação epitelial irregular;
- Perda de polaridade das células basais;
- Cumes rete em forma de gota;
- Aumento do número de imagens mitóticas;
- Mitoses anormalmente superficiais;
- Queratinização prematura em células únicas (disqueratose);
- Pérolas de queratina dentro de cumes rete;

Essa lesão normalmente é categorizada em três níveis de evolução: leve, moderado e grave. Os níveis variam de acordo com a gravidade de cada um dos critérios mencionados, quanto maior a gravidade dos critérios, maior será a severidade da displasia no tecido avaliado. Na Figura 2 são apresentados exemplos de tecidos histológicos da cavidade oral com presença de displasia.

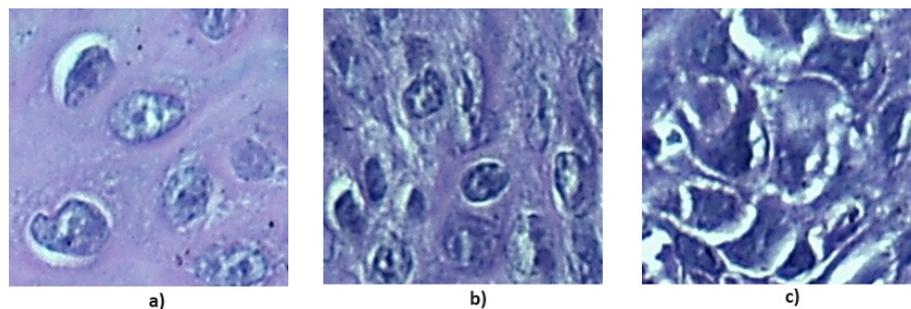


Figura 2 – Exemplos de tecido com displasia: a) representa um tecido com displasia leve, b) um exemplo de displasia moderada e c) amostra de displasia severa.

Os sistemas de diagnóstico auxiliado por computador (do inglês, *Computer-Aided Diagnosis - CAD*) são construídos com o intuito de fornecer uma segunda opinião aos profissionais da saúde com o intuito de ajudar a identificar padrões, anomalias ou características relevantes nos exames que são dificilmente perceptíveis a olho humano. Por meio da aplicação de algoritmos computacionais, o sistema CAD pode destacar áreas suspeitas, coletar informações importantes e então fornecer dados ao especialista nas etapas de um diagnóstico.

2.3 Sistema de Diagnóstico Auxiliado por Computador

O CAD é uma ferramenta que utiliza tecnologias computacionais para auxiliar os profissionais de saúde no processo de diagnóstico de doenças. Essa abordagem é uma combinação de dados clínicos e informações providas por meio de exames como, imagens de raios-X, tomografia computadorizada e outros. Esse método utiliza algoritmos complexos e inteligência artificial para auxiliar na análise e interpretação das imagens (HALLI; MAKANDAR, 2017). Esse processo pode agilizar o diagnóstico, permitindo um tratamento mais rápido e eficaz para os pacientes. O sistema geralmente é dividido em algumas etapas, como demonstrado na Figura 3.



Figura 3 – Etapas empregadas em um sistema CAD para auxiliar um especialista.

A etapa de pré-processamento possui como objetivo fazer o melhoramento das

imagens, de modo a fazer com que as próximas etapas tenham um melhor desempenho na detecção e análise de cada uma delas. Nessa etapa pode-se aplicar, por exemplo, uma normalização dos corantes H&E nas imagens de forma a tentar diminuir a variância causada pela diferença de iluminação e problemas com os corantes durante o processo de coloração. A etapa de segmentação tem o objetivo de separar regiões das imagens que são de maior interesse para realizar a identificação da lesão. Nessa etapa, pode-se aplicar, por exemplo, algoritmos que fazem uso da diferença entre a intensidade dos pixels para diferenciar objetos de fundo como os algoritmos baseados em limiarização ou detecção de bordas (YU et al., 2023).

A etapa de extração de características consiste em transformar as regiões separadas na etapa de segmentação em características numéricas que descrevem propriedades nas imagens. Há vários tipos de descritores que podem ser extraídos e analisados, dentre eles: i) atributos morfológicos, que se baseiam em formato e tamanho dos objetos; ii) atributos de cores e intensidade de brilho; iii) atributos de textura (SOILLE, 2002), dos quais se destaca a matriz de concorrência (SFERRAZZA, 2023) e a geometria fractal (MANDELBROT, 1975).

Por fim, a última etapa do CAD é a classificação, em que ocorre o reconhecimento dos padrões a partir dos atributos. Para isso, é aplicado um algoritmo classificador que recebe as características extraídas de cada imagem e procura por padrões nesses conjuntos. Para reconhecer os padrões, o algoritmo geralmente é dividido em estágios de treinamento e teste. No treinamento, o classificador é modelado com um conjunto de características pré-rotuladas. No estágio de teste o classificador atua sobre um conjunto de imagens não rotuladas para avaliar o seu desempenho sobre essas imagens. Vários algoritmos de classificação estão disponíveis na literatura e são amplamente utilizados, dentre eles, se destacam as redes neurais (YAMASHITA et al., 2018), árvore de decisão (ROKACH; MAIMON, 2005) e floresta aleatória (BREIMAN, 2001). Os descritores de textura são fundamentais para auxiliar na captura de informações sobre a distribuição espacial de padrões em uma imagem, permitindo a caracterização das texturas presentes nas regiões analisadas, as quais são fundamentais para a classificação precisa das imagens.

2.4 Análise de Textura

Durante a etapa de extração de características, os dados obtidos pelos descritores são adicionados ao vetor de características para aplicação sobre os algoritmos de classificação para uma análise das propriedades em comuns à classe de padrões. Há diversas técnicas de descritores que exploram informações a serem analisadas, em que as principais são categorizadas em: descritores morfológicos, são também denominados de descritores de fronteira, e descritores não-morfológicos, também chamados de descritores

regionais (SOILLE, 2002).

Os descritores morfológicos simbolizam as margens e bordas das imagens. Por outro lado, os descritores não morfológicos retratam as características internas, por exemplo, os pixels que formam as regiões internas das estruturas. As características não morfológicas são vistas como problemáticas devido às dificuldades de interpretação das mudanças das informações internas do olho humano. Uma classe relevante para extração dessas informações pode ser obtida pelas propriedades de textura da imagem. Há uma vasta quantidade de modos de processamento de imagem para essa operação. Dentre eles, tem-se as funções de autocorrelação, métodos apoiados em modelos, matriz de coocorrência, geometria fractal (MEDJAHED, 2015). Em meio às abordagens, as técnicas com conceito estatístico, como matriz de coocorrência, retratam indiretamente a textura por meio de propriedades não determinísticas que definem distribuições e relacionamentos entre os níveis de cinza ou cores dos pixels de uma imagem. Utilizando essa abordagem, é possível qualificar a textura como áspera, suave, granulada, entre outros. Para esse trabalho, a análise de textura será feita a partir do uso de descritores baseados em geometria fractal.

2.5 Geometria Fractal

Os fractais são formas geométricas que apresentam semelhança consigo mesmas em diferentes escalas. Isso implica que, ao examinar uma porção específica de um fractal, sua configuração se assemelha à forma global. Essa característica de auto-semelhança persiste, indefinidamente, conforme a escala é aumentada ou reduzida (ASSIS et al., 2008). A obtenção de descritores fractais envolve o uso de equações recursivas, as quais são frequentemente realizadas por meio de algoritmos computacionais. Um exemplo clássico conhecido de fractal é o conjunto de *Mandelbrot* (RUBIANO, 1996), gerado iterativamente pela aplicação de uma equação matemática simples a cada ponto em um plano complexo.

Além de sua atratividade estética, os fractais possuem aplicações em diversas áreas como, modelagem de fenômenos naturais, simulações computacionais, arte digital e na compreensão de padrões em sistemas complexos. Eles representam uma expressão visual fascinante da interligação entre matemática, ciência e arte. Pode existir certo nível de complexidade no que se refere à construção de programas que façam identificação (descrição) de objetos, uma vez que eles podem aparecer em diversas formas que variam de acordo com a natureza como, por exemplo, relâmpagos e nuvens, esses não podem ser identificados com a utilização de estruturas geométricas euclidianas simples como círculo, triângulo, polígono etc. Contudo, o conceito da geometria fractal permite um melhor estudo sobre essas formas (MANDELBROT; FRAME, 1997). Para as estruturas irregulares pode-se criar algumas regras de acordo com os modos em que o objeto pode ser observado,

por exemplo, regras que relacionam à escala de observação do objeto, como mostrado na Figura 4.

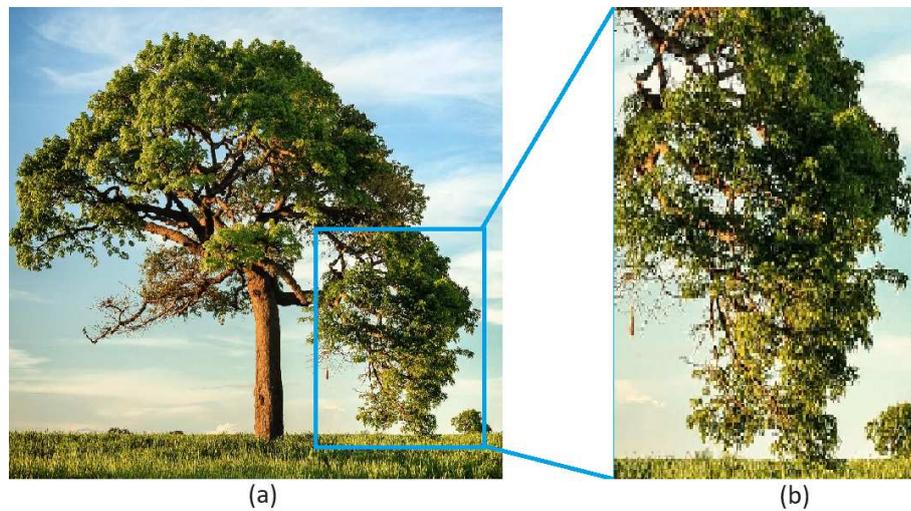


Figura 4 – Imagem de uma árvore sob duas escalas de observação diferentes: imagem completa (a) e imagem observada sob uma escala 3 vezes menor (b).

É possível caracterizar um fractal enumerando alguns pontos (FLORINDO, 2013), sendo os principais deles:

- Dificuldade em descrever o objeto através da geometria convencional;
- Autossimilaridade aproximada.

Para extrair descritores baseados em geometria fractal, alguns algoritmos estão presentes na literatura, entre eles, destaca-se o algoritmo *gliding-box* (IVANOVICI; RICHARD; DECEAN, 2009) e o algoritmo *box-counting* (NIKOLAIDIS; NIKOLAIDIS; TSOUROU, 2011). Ambos algoritmos exploram as propriedades de autossimilaridade, nos quais as imagens são separadas em escalas diferentes e, para cada imagem, características são extraídas e novos atributos são criados.

O algoritmo *gliding-box* envolve posicionar uma caixa de dimensões $L \times L$ no canto superior esquerdo da imagem, onde L representa o lado da caixa em pixels. A caixa é movida da esquerda para a direita até a região inferior da imagem, percorrendo todos os pixels. Após passar por toda a imagem, a caixa é reposicionada no ponto inicial, e o valor de L é aumentado em 2 unidades, conforme ilustrado na Figura 5. É crucial que L seja sempre um número ímpar, pois a existência de um pixel central é necessária para as etapas subsequentes do método. Em uma imagem de tamanho $H \times W$, o número total T de caixas B_i em relação a L é expresso pela Equação 2.1.

$$T(L) = (H - L + 1) \times (W - L + 1) \quad | \quad L \leq \min(H, W). \quad (2.1)$$

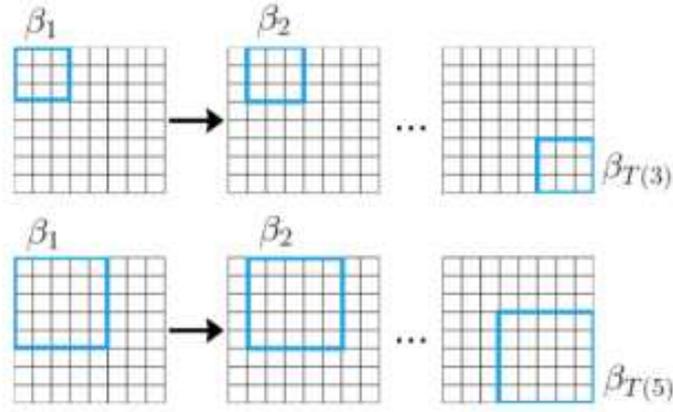


Figura 5 – Simulação do algoritmo *gliding-box* em uma imagem com dimensões 8×8 .
Fonte: (ROBERTO et al., 2021)

Procede-se a uma análise de semelhança de cor para cada pixel contido em β_i , sempre que uma caixa β_i é movida. Essa análise é conduzida ao fixar o pixel central e atribuir-lhe um vetor $f_c = r_c, g_c, b_c$ que representa as intensidades luminosas nos diferentes canais de cores, considerando imagens no sistema de cores *Red*, *Green* e *Blue* (RGB). Os demais pixels presentes na caixa são associados a um segundo vetor $f_i = r_i, g_i, b_i$ e são comparados ao pixel central por meio do cálculo de uma distância de cor Δ . Esse procedimento possibilita identificar quais pixels pertencem ao hiperespaço RGB formado pelo centro da caixa β_i (IVANOVICI; RICHARD; DECEAN, 2009). As distâncias *Chessboard* (Δ_h), *Euclidean* (Δ_e) e *Manhattan* (Δ_m), seguindo as Equações 2.2, 2.3, 2.4 podem ser usadas para Δ .

$$\Delta_h = \max(|f_i(k_i) - f_c(k_c)|), k \in r, g, b. \quad (2.2)$$

$$\Delta_e = \sqrt{\sum_{k=r, g, b} (f_i(k_i) - f_c(k_c))^2}. \quad (2.3)$$

$$\Delta_m = \sum_{k=r, g, b} |f_i(k_i) - f_c(k_c)|. \quad (2.4)$$

Caso o valor de Δ seja menor ou igual que a escala L , então assume-se que f_i tem valor 1, caso contrário, tem valor 0. Na Figura 6, é representada a forma como os valores são atribuídos em uma caixa com escala $L = 3$.

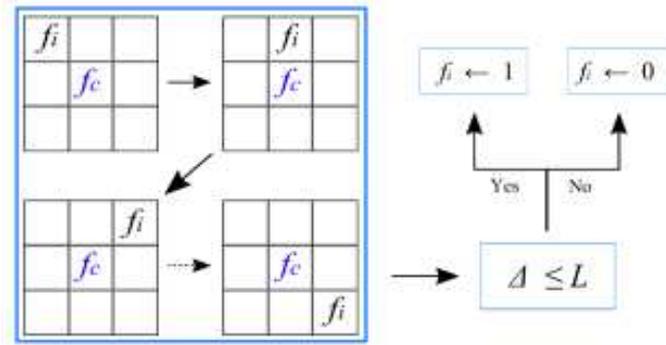


Figura 6 – Atribuindo os valores aos pixels em uma caixa 3×3 depois de calcular o valor da distância Δ . Fonte: (ROBERTO et al., 2021)

Após calcular a quantidade de pixels que satisfazem Δ para diferentes valores de L , a informação obtida é utilizada para construir uma matriz de probabilidade $P(m,L)$. Os valores da matriz representam a probabilidade P de m pixels serem atribuídos ao valor 1, utilizando uma escala L . Por fim, a matriz é normalizada de forma que a soma de todos os valores em uma coluna seja igual a 1. Depois que toda a matriz de probabilidade for preenchida, os atributos fractais locais podem ser obtidos. Há diferentes tipos de atributos que podem ser extraídos com a geometria fractal, destacam-se: a dimensão fractal, a lacunaridade e a percolação (LOPES; BETROUNI, 2009).

2.5.1 Dimensão Fractal

A dimensão fractal (do inglês, *fractal dimension* - FD) é a métrica mais comum para acessar atributos fractais em uma imagem. Em 1961, houve a primeira pesquisa sobre o assunto, em que Richardson tentou medir o comprimento do litoral da Grã-Bretanha. Ele observou que quanto menor a escala de medida, maior era o comprimento do litoral (RICHARDSON, 1960) e chegou a conclusão de que isso acontecia, pois, em menores escalas, mais detalhes como, pequenas enseadas, riachos ou penínsulas, contribuíam para a soma do comprimento final da estrutura analisada. Em 1983, Mandelbrot observou um comportamento similar nos fractais (MANDELBROT, 2015). Ainda, foi observado que uma linha reta é obtida se o logaritmo da medida escalar foi projetada em um gráfico que relaciona também o logaritmo do comprimento total. O valor do coeficiente angular dessa linha reta permite definir a dimensão fractal do objeto, e os valores individuais para cada escala de medida correspondem às dimensões fractais locais. Em um modelo inicial, a dimensão fractal foi calculada apenas para os objetos com até 2 dimensões topológicas ou imagens binárias.

Com o desenvolvimento de novas abordagens, foi possível calcular o FD para escalas de cinza ou imagens coloridas. De acordo com os autores em Ivanovici e Richard (2010), características locais (denominadas $D(L)$) da dimensão fractal podem ser obtidas

de acordo com a criação de uma matriz probabilística definida pela Equação 2.5.

$$D(L) = \sum_{m=1}^{L^2} \frac{P(m,L)}{m}. \quad (2.5)$$

Quando diferentes imagens possuem valores similares ou, até mesmo idênticos, pode ser um problema, como exemplificado na Figura 7 nos *Sierpinski* fractais. Contudo, esse problema pode ser resolvido com o uso de outras medidas baseadas em geometria fractal, como a lacunaridade.

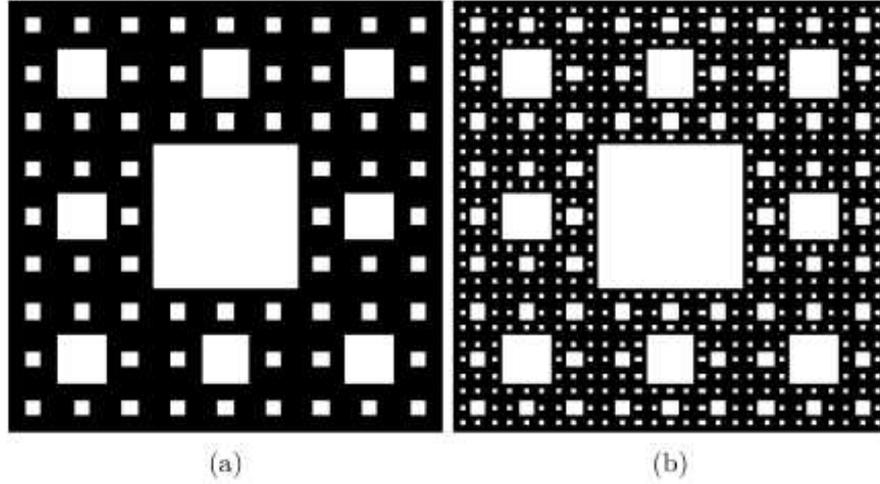


Figura 7 – Tapetes de *Sierpinski*. As imagens (a) e (b) possuem similares valores FD, contudo eles possuem valores diferentes de lacunaridade. Fonte: (MANDELBROT, 2015)

2.5.2 Lacunaridade

A lacunaridade (do inglês, *lacunarity* - LAC) é uma medida complementar a FD que permite avaliar como o espaço de um fractal é preenchido. Como na medida FD, os autores em (IVANOVICI; RICHARD; DECEAN, 2009) também propuseram um método para obter a LAC em imagens coloridas. Esse método tem passos similares ao método de obtenção FD para construção da matriz probabilística. Então, o primeiro e o segundo momentos são calculados usando as Equações 2.6 e 2.7. Os valores m e $P(m,L)$ são obtidos do mesmo modo como descritos para FD, na equação 2.5. O descritor LAC mede em L escala dada por $\Lambda(L)$, usando a Equação 2.8.

$$\mu(L) = \sum_{m=1}^{L^2} mP(m,L). \quad (2.6)$$

$$\mu^2(L) = \sum_{m=1}^{L^2} m^2P(m,L). \quad (2.7)$$

$$\Lambda(L) = \frac{\mu^2(L) - (\mu(L))^2}{(\mu(L))^2}. \quad (2.8)$$

2.5.3 Percolação

A percolação (do inglês, *percolation* - PERC) é um fenômeno físico descrito por [Broadbent e Hammersley \(1957\)](#), que se refere ao movimento de fluidos através de meios porosos. Diversos sistemas naturais manifestam esse comportamento, como a passagem de água pelo pó de café ou de gases por meio de uma substância rochosa. Quando um desses fluidos, seja água ou gás, consegue atravessar todo o sistema de uma extremidade à outra, ocorre o que é chamado de fenômeno de percolação. Esses conceitos também podem ser aplicados à análise de imagens. A Figura 8 ilustra um exemplo de uma grade com dimensões de 15x15 pixels, demonstrando como a percolação acontece. Nesse exemplo, alguns pixels da grade foram marcados com o símbolo (*). Em uma analogia com a teoria da percolação, esses pixels marcados representam poros, que são espaços pelos quais um suposto fluido pode atravessar o meio. Nota-se que alguns pixels formam aglomerados quando se aplica uma vizinhança de 4 pixels, como mostrado na Figura 8(b).

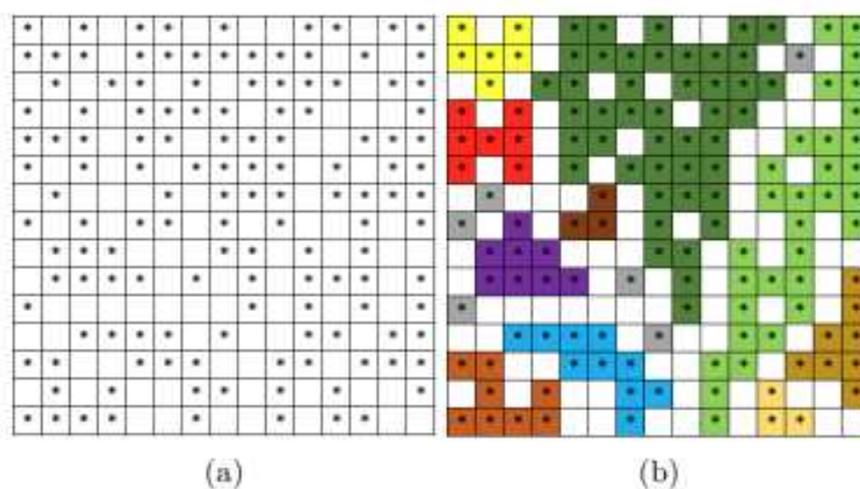


Figura 8 – Uma matriz com dimensões de 15×15 , em (a) a percolação se manifesta devido à presença de um aglomerado que se estende desde uma extremidade até a outra da matriz, sendo visualizado em uma tonalidade de verde claro (b). Fonte: ([ROBERTO et al., 2021](#)).

Para derivar atributos de percolação de uma imagem, por meio da aplicação de conceitos da geometria fractal, surgiu uma proposta em 2017 ([ROBERTO et al., 2017](#)). Esse método se fundamenta nos Modelos de ([IVANOVICI; RICHARD; DECEAN, 2009](#)), os quais possibilitam a obtenção de Dimensão Fractal (FD) e Comprimento de Área Lacunar (LAC) a partir de imagens coloridas ([IVANOVICI; RICHARD; DECEAN, 2009](#)). A abordagem segue as etapas do algoritmo *gliding-box* e do cálculo da distância de *Minkowski* entre um pixel central e os demais pixels contidos em uma caixa, visando analisar a similaridade dos pixels coloridos. Posteriormente, os pixels que atendem ao critério da distância de *Minkowski*, definidos como poros, são identificados por meio do algoritmo

de rotulagem *Hoshen-Kopelman* (HOSHEN; BERRY; MINSER, 1997). São extraídas três características distintas para cada escala L : o número médio de aglomerados por compartimento, a taxa de ocorrência de percolação e a área média do maior aglomerado. O número médio de aglomerados $C(L)$ é calculado pela Equação 2.9, que envolve a soma do número total de aglomerados c em cada caixa i , dividida pelo número total de caixas T .

$$C(L) = \frac{\sum_{i=1}^{T(L)} c_i}{T(L)}. \quad (2.9)$$

A Equação 2.10 define o cálculo da taxa de ocorrência de percolação $Q(L)$.

$$q_i = \begin{cases} 1, & \frac{\Omega}{L^2} \geq 0.59275. \\ 0, & \frac{\Omega}{L^2} < 0.59275. \end{cases}$$

$$Q(L) = \frac{\sum_{i=1}^{T(L)} q_i}{T(L)}. \quad (2.10)$$

A expressão na Equação 2.11 representa a computação da área média de cobertura dos maiores aglomerados, denotada como $M(L)$. Esse valor é obtido somando as taxas de ocupação dos maiores aglomerados em cada caixa M_i e dividindo pelo número total de caixas L^2 .

$$M(L) = \frac{\sum_{i=1}^{T(L)} \frac{M_i}{L^2}}{T(L)}. \quad (2.11)$$

A quantidade de características locais adquiridas é influenciada pela escala total de observação L . Quando L varia de 3 a L_{max} com um acréscimo de 2 unidades, a quantidade de recursos locais equivale a $5 \times \left(\frac{L_{max}-3}{2} + 1\right)$ para cada incremento Δ .

2.6 Classificação

A classificação é baseada em algoritmos de aprendizado de máquina, e é usado para atribuir um rótulo de classe a uma entrada de dados. O algoritmo pode receber como entrada uma imagem radiográfica de um pulmão e a saída do algoritmo será um rótulo para a imagem, a depender do objetivo, a saída poderia ser o estado de saúde do paciente: "Pulmão saudável", "Pulmão com indícios de Covid-19", etc.

Existem dois tipos de classificação dos dados, a classificação supervisionada e a não-supervisionada. Na supervisionada todas as imagens do banco de dados estão previamente rotuladas, ou seja, é possível saber qual a saída correta esperada para cada imagem do banco. Na abordagem não-supervisionada as imagens não possuem rótulos pré-definidos, portanto, cabe ao algoritmo encontrar os padrões nesse conjunto de imagens (ALZUBAIDI et al., 2021). Diversos algoritmos são propostos para classificação de dados, tais como:

regressão linear, regressão logística, *naive bayes*, árvore de decisão, floresta aleatória e vários outros.

Para este trabalho, alguns algoritmos de redes convolucionais neurais foram usados, tais como, MobileNet, ResNet50 e ResNet50V2, além disso, algoritmos de *machine learning* também foram investigados, por exemplo, árvore de decisão, floresta aleatória e regressão logística. Esses algoritmos apresentaram resultados interessantes em diversos trabalhos da literatura, entre eles, (SILVA et al., 2023) e (ALJAME et al., 2020), por esse motivo esses algoritmos foram escolhidos para serem estudados neste trabalho.

2.6.1 Redes Neurais Convolucionais

As redes neurais convolucionais (do inglês, *convolutional neural network* - CNN) são estruturas de rede criadas para aprender categorias espaciais de recursos de forma automática e adaptativa, a implementação consiste em vários blocos de construção, como camadas de convolução, camadas de *pooling*, camadas *flatten* e camadas totalmente conectadas (YAMASHITA et al., 2018).

As camadas de convolução dispõem de vários filtros com parâmetros que devem ser aprendidos. As proporções desses filtros precisam ser menores que a da entrada, pois a máscara do filtro é movida sobre a entrada e é feito um produto escalar entre os pesos do filtro do *kernel* e os valores da imagem em cada parcela da entrada, produzindo como saída um mapa de ativação 2D. A Figura 9 mostra um exemplo desse processo aplicado a uma imagem de entrada, em que a submatriz em cor laranja na entrada representa a parcela que está a ser multiplicada pelo filtro. Essa multiplicação tem como resultado final a célula destacada em laranja no mapa de ativação. O mesmo raciocínio se aplica para o cálculo das posições restantes do mapa (ALZUBAIDI et al., 2021) (YAMASHITA et al., 2018).

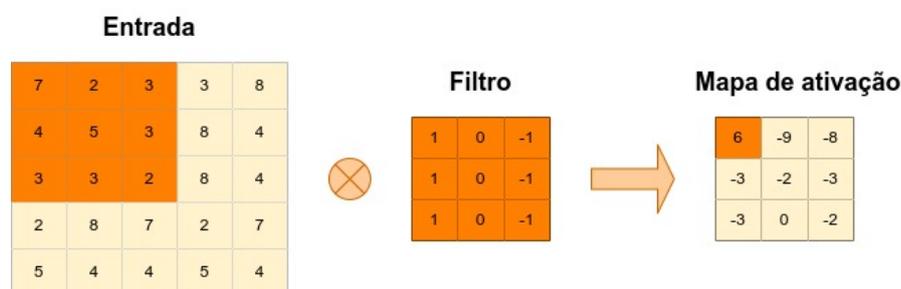


Figura 9 – Exemplo de como ocorre uma convolução.

O método de aplicação da convolução envolve o conceito de alguns parâmetros:

- **A quantidade de filtros:** Esse valor determina a quantidade de canais resultantes na saída, pois cada um dos filtros usados resulta em um mapa de ativação

2D (SANTOS et al., 2022).

- **O tamanho dos filtros:** Corresponde às grandezas (altura x largura) de cada filtro (SANTOS et al., 2022).
- **stride:** É a quantidade de pixels que serão investigados e transformados a cada passo. Esse valor é inversamente proporcional ao tamanho da saída, quanto maior é esse valor, menor será a grandeza de saída (SANTOS et al., 2022).
- **padding:** Define se há a necessidade de que a entrada seja automaticamente preenchida com pixels zeros nas bordas, desse modo é possível garantir que os filtros processem todos os valores de entrada. (SANTOS et al., 2022).

As camadas de *pooling* são encarregadas de reduzir os parâmetros dos dados de entrada. Nessas camadas, existem várias abordagens diferentes, a mais utilizada é o *max pooling* que consiste em separar a entrada em conjuntos de mesmo tamanho e cada um desses conjuntos será representado pelo máximo valor contido. Outra abordagem de *pooling* que pode ser aplicada é o *average pooling* que segue a mesma lógica do anterior, separando a entrada em conjuntos de mesmo tamanho. Entretanto, nesta última abordagem, cada um dos conjuntos será representado pela média de todos os valores inclusos. A Figura 10 ilustra os processos de *max pooling* e *average pooling* em uma entrada utilizando $stride = 2$, onde a célula em amarelo na matriz de *max pooling* representa o máximo de todos os valores da submatriz que está em amarelo na entrada. Na matriz de *average pooling*, a célula em amarelo se refere a média dos valores da submatriz. O mesmo raciocínio se aplica para as outras cores em todas as matrizes apresentadas (YAMASHITA et al., 2018).

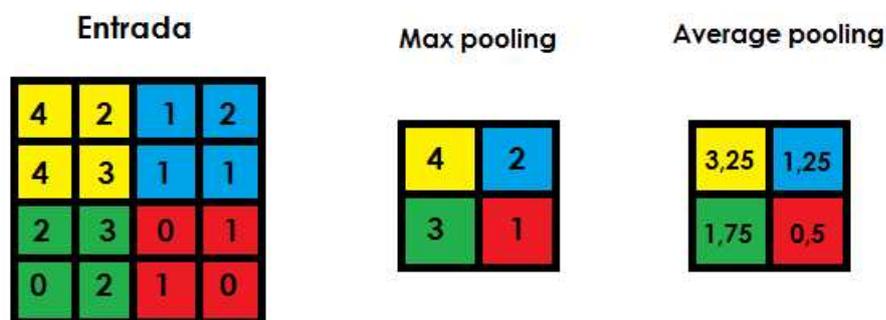


Figura 10 – Exemplo de abordagens de *max pooling* e *average pooling*.

A camada *flatten* tem o objetivo de reduzir um *array* multidimensional em um *array* unidimensional. Essa abordagem é geralmente usada como transição para uma camada totalmente conectada. Na Figura 11 é exemplificado o processo nessa camada, recebendo como entrada uma matriz com dimensões (3,3). A camada teve como saída um

array unidimensional com 9 posições (ALZUBAIDI et al., 2021) (YAMASHITA et al., 2018).

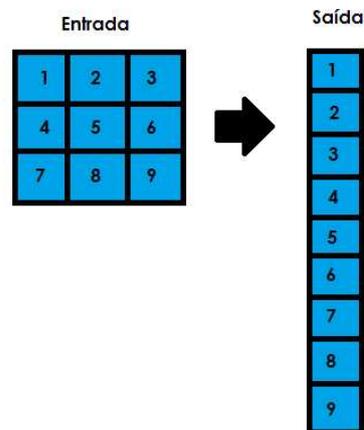


Figura 11 – Demonstração de uma transformação feita na camada *flatten*.

As camadas totalmente conectadas normalmente são as últimas da rede. Nessas camadas, as entradas estão diretamente ligadas a cada unidade de ativação da próxima camada. A primeira camada fortemente conectada tem o objetivo de configurar os pesos das arestas de modo que as predições corretas sejam as melhores possíveis. A última camada totalmente conectada é responsável por atribuir uma probabilidade final para cada classe indicando a probabilidade da imagem de entrada pertencer a cada uma das classes. (ALZUBAIDI et al., 2021) (YAMASHITA et al., 2018).

Há vários modelos de *CNN*, e eles se diferenciam pela sua estrutura, mas possuem os mesmos objetivos gerais, que é fazer a extração de características de um conjunto de dados, treinar a rede com características extraídas e fazer a classificação dos dados. Como exemplo de rede neural, existe a *AlexNet*, uma *CNN* clássica com uma estrutura composta por 8 camadas, dentre elas, as 5 primeiras são camadas de convolução, algumas seguidas por camadas de *max pooling*, e, por fim, 3 camadas totalmente conectadas. A Figura 12 ilustra um modelo *AlexNet*.

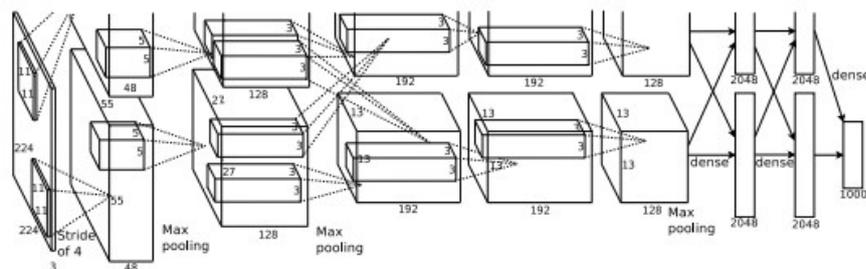


Figura 12 – Estrutura da arquitetura *AlexNet* Fonte: (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

Por meio da investigação de alguns modelos *CNNs*, principalmente a *ResNet*, serão estudadas as características resultantes dessas arquiteturas com o propósito de representar as informações das lesões nas imagens histológicas.

2.6.2 Árvore de Decisão

Na computação uma das principais estruturas de dados é a árvore, a qual possui vários nós e cada um deles guarda informações. O primeiro nó da árvore é denominado raiz, esse é o nó que está acima de todos os outros. Cada um dos nós fazem ligações com outros, que são denominados filhos, pois estão em posição hierárquica inferior. Por último, existe o nó folha que é o nó que não tem filhos e guardam informações finais. De forma análoga, o método da árvore de decisão funciona de forma que cada nó da árvore guarda uma regra, o nó subsequente dependerá da resposta do nó atual, ou seja, se a regra é satisfeita ou não. À medida que se caminha na árvore as decisões são tomadas e podem levar, por último, a um nó folha, que representa a decisão final a ser efetuada (TOPÎRCEANU; GROSSECK, 2017). A Figura 13 é um exemplo de uma árvore de decisão que faz uma predição sobre três variáveis (x_1 , x_2 , x_3) acerca de três constantes (A , B , C) em um determinado problema. Nesse exemplo, os vértices em verde são considerados folhas e, portanto, contém informações finais. Se, por exemplo, o caminho na árvore atingir o vértice 6, então pode-se afirmar que $x_1 < A$ e $x_2 < B$, devido às informações das arestas percorridas no caminho. Por outro lado, atingir o vértice 3, então pode-se afirmar que $x_1 \geq A$ e $x_3 < C$. A mesma lógica se aplica aos outros possíveis caminhos.

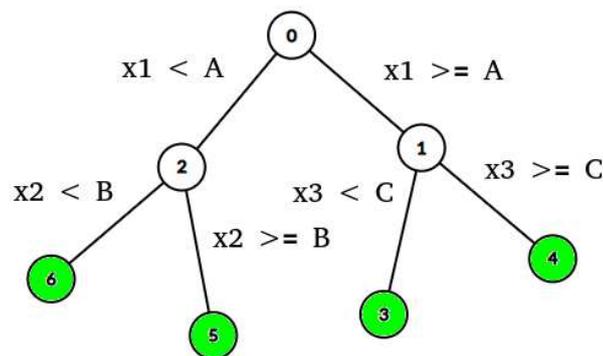


Figura 13 – Exemplo de um Árvore de decisão.

2.6.3 Floresta Aleatória

O algoritmo de classificação floresta aleatória (do inglês, *random forest* - RF) é um algoritmo de aprendizagem supervisionada. Ele se dá pela combinação de diversas árvores de decisões, de modo a obter resultados com maior acurácia. O algoritmo começa selecionando vários conjuntos de características aleatórias do banco de dados, após isso,

constrói uma árvore de decisão para cada conjunto selecionado e realiza o treinamento. Após essa etapa, haverá várias árvores de decisão treinadas com diferentes características. Com isso, para fazer a predição de uma determinada imagem, o algoritmo recebe como entrada as características extraídas da imagem e realiza uma predição em cada uma das árvores de decisão criadas no treinamento sobre a entrada. Cada árvore da floresta produz uma predição sobre as características de entrada. A predição final da floresta será a classe que foi resultante do maior número de árvores de decisões no processo (SARICA; CERASA; QUATTRONE, 2017). A Figura 14 ilustra como o processo de classificação ocorre sobre uma entrada qualquer, considerando uma floresta já treinada sobre um conjunto de descritores. Nesse exemplo, a floresta possui 3 árvores de decisões, cada uma delas possui configuração diferente, pois foram treinadas com diferentes características. Cada uma das árvores representa a predição resultante sobre uma mesma entrada. A primeira árvore teve como saída a classe 0, a segunda teve como saída a classe 1 e, por fim, a terceira teve como saída a classe 0. Portanto, como a classe 0 apareceu mais vezes, o resultado da predição da floresta será a classe 0. Devido ao grande número de árvores de decisões empregadas no processo, o método possui elevada acurácia e também é robusto, pois cada árvore de decisão funciona em tempo logarítmico.

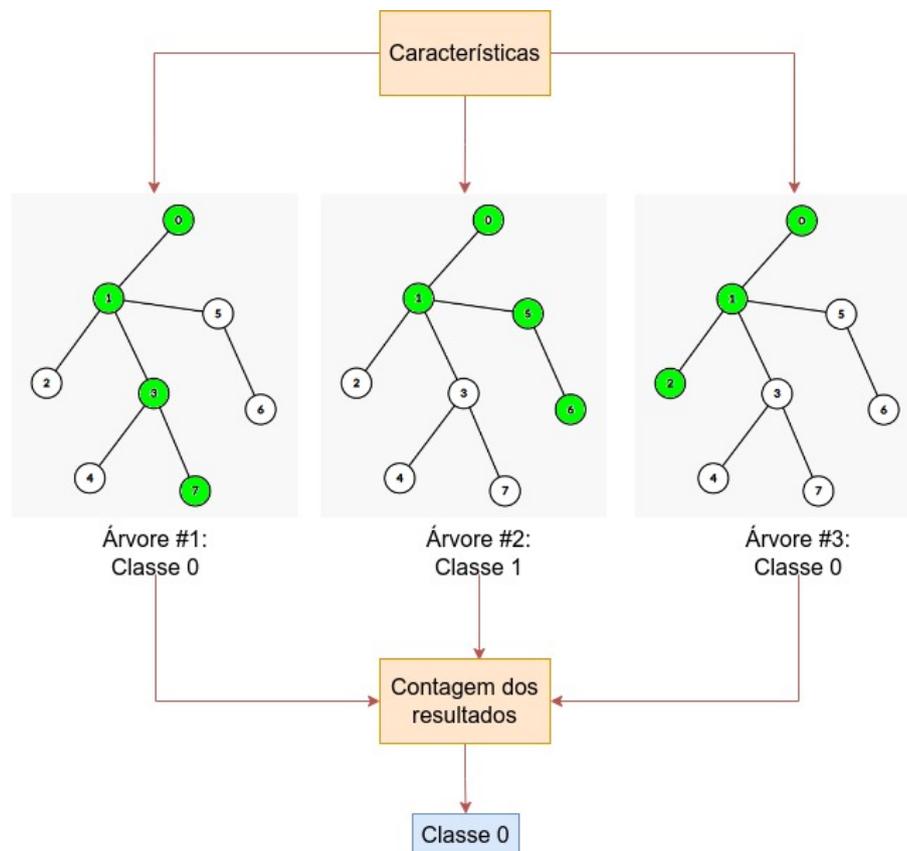


Figura 14 – Exemplo de uma floresta aleatória.

2.6.4 Regressão Logística

A Regressão Logística (do inglês, *logistic regression*(LR)) é um algoritmo de aprendizagem supervisionada que produz um modelo que permite a predição de valores tomados por uma variável categórica. Esse algoritmo visa estimar a probabilidade de ocorrência de um evento, como um voto, utilizando um conjunto específico de dados de variáveis independentes. Nesse método, é aplicada uma transformação *logit* às probabilidades, isto é, a razão entre a probabilidade de sucesso e a probabilidade de falha. Essa transformação, também conhecida como *log odds* ou logaritmo natural das probabilidades, é expressa nas equações 2.12 e 2.13.

$$\text{logit}(p_i) = \frac{1}{1 + e^{-p_i}}. \quad (2.12)$$

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \sum_{i=1}^k (\beta_i * x_i). \quad (2.13)$$

Na equação 2.12, a variável dependente ou de resposta, denotada por $\text{logit}(p_i)$, é influenciada pela variável independente x . Dentro do escopo deste modelo, o parâmetro β , também referido como coeficiente, é comumente estimado por meio da técnica de Estimaco de Verossimilhana Mxima (MLE) (SOUSA, 2018). Este mtodo explora uma variedade de valores para beta por meio de iteraes, visando otimizar o ajuste dos *log odds* (POHL et al., 2007). Cada iterao gera a funo *log likelihood* (ETZ, 2018), sendo o objetivo da regresso logstica maximizar essa funo para obter a estimativa ideal do parmetro.

Uma vez identificado o coeficiente ideal (ou coeficientes, em casos de mltiplas variveis independentes), torna-se possvel calcular as probabilidades condicionais para cada observao. Tais probabilidades so registradas e somadas para gerar uma probabilidade prevista. No mbito da classificao binria, uma probabilidade inferior a 0,5 prev o resultado 0, enquanto uma probabilidade superior a 0 indica a previso do resultado 1.

Aps o clculo do modelo, uma prtica recomendada  avaliar quo eficazmente ele prev a varivel dependente, medida conhecida como qualidade de ajuste. O teste de *Hosmer-Lemeshow* (PAUL; PENNELL; LEMESHOW, 2013) destaca-se como uma abordagem popular para avaliar o ajuste do modelo.

2.7 Aumento de dados

O aumento de dados  uma tcnica que consiste em ampliar o banco de imagens criando novos exemplares que derivam das imagens j existentes com o objetivo de ter ganhos no aprendizado. Essa abordagem pode ser muito vantajosa no que se refere aos

problemas de classificação de imagens, quando a quantidade de dados disponíveis é limitada ou desbalanceada, nesses casos, pode-se usar dessa técnica para aumentar essa quantidade e diversificar as amostras (PEREZ; WANG, 2017).

Há várias técnicas de aumento de dados que são amplamente utilizadas, algumas delas são:

- Corte aleatório: Consiste em cortar uma parte aleatória da imagem;
- Espelhamento horizontal: Consiste em virar a imagem no eixo x;
- Espelhamento vertical: Consiste em virar a imagem no eixo y;
- Rotação: Consiste em girar uma imagem em um ângulo definido;
- Translação: Consiste em deslocar a imagem em uma direção;
- Zoom: Consiste em aumentar/reduzir a escala da imagem;

Essas técnicas podem ajudar a prevenir *overfitting* em modelos, além de ajudar a aumentar a acurácia (MUMUNI; MUMUNI, 2022). Na Figura 15 é apresentada a técnica de aumento baseada na operação de rotação de uma imagem.



Figura 15 – Exemplo de aumento de dados utilizando a técnica de rotação em 90 graus: imagem original (a) e imagem rotacionada (b).

2.8 Regularizadores

Quando se trabalha com modelos baseados em aprendizado de máquina, é preciso ter ciência de que o aprendizado nem sempre tende a ser a melhor e pode levar a péssimos resultados. Como ilustração disso, há um fenômeno chamado sobre-ajuste (do inglês,

overfitting) que ocorre quando um modelo de aprendizado de máquina faz um bom desempenho nos conjuntos de dados treinados, mas acaba não realizando uma performance satisfatória em dados diferentes dos que participaram da fase de aprendizado (WONG et al., 2018). Há algumas abordagens eficientes de regularização, tais como: *L1*, *L2* e *dropout*. A regularização *L1* tem o objetivo de diminuir a variância do modelo e diminuir a penalização da seleção de características com muita correlação, tentando selecionar apenas uma característica e zerando o coeficiente das demais.

Na regularização *L1*, adiciona-se o seguinte valor *L1* à função de custo da rede (Equação 2.14).

$$L1 = \lambda \sum_{i=1}^N (|w_i|). \quad (2.14)$$

Após adicionar *L1* à função de custo do modelo $L(x,y)$ (Equação 2.15), tem-se um modelo de rede com regularização *L1* aplicada (Equação 2.16).

$$L(x,y) = \sum_{i=1}^N (y_i - h_w(x_i))^2, \quad (2.15)$$

$$L1(x,y) = \sum_{i=1}^N (y_i - h_w(x_i))^2 + L1. \quad (2.16)$$

Na Equação 2.14, a variável w é o parâmetro a ser multiplicado pela entrada do modelo para conseguir uma predição. O valor de λ será usado para penalizar os parâmetros da rede e deve ser pré-definido antes de iniciar o treinamento. A função $h_w(x_i)$ é o resultado da predição da entrada x_i . Nessa abordagem, tem-se a desvantagem de que os parâmetros acabam ficando muito esparsos, os valores dos parâmetros ficam próximos de zero.

Na regularização *L2*, a ideia é próxima a abordagem *L1*, contudo o parâmetro w é elevado ao quadrado (Equação 2.17).

$$L2 = \lambda \sum_{i=1}^N w_i^2. \quad (2.17)$$

Adicionando esse valor *L2* à função de custo do modelo $L(x,y)$, tem-se a regularização *L2* aplicada (Equação 2.18). Uma grande diferença do método *L1* para o *L2* é que no *L2* os resultados não são tão esparsos quanto no *L1*, devido aos valores serem elevados ao quadrado.

$$L2(x,y) = \sum_{i=1}^N (y_i - h_w(x_i))^2 + L2. \quad (2.18)$$

A regularização *dropout*, muitas vezes já se encontra aplicada por padrão em alguns modelos. Parte dos neurônios da rede são escolhidos aleatoriamente e são desativados de acordo com uma probabilidade (p) a ser definida. Nessa abordagem, os nós da rede não serão atribuídos com valores muito altos, o que faz com que esses valores sejam

mais dispersamente distribuídos. Assim, a saída do modelo não dependerá apenas de um nó (TIAN; ZHANG, 2022). A Figura 16 mostra um exemplo de como ocorre a execução da camada *dropout* em um modelo.

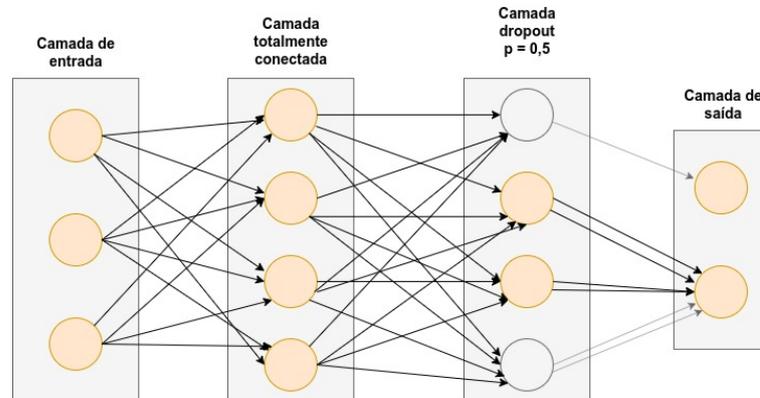


Figura 16 – Ilustração de uma etapa de um modelo usando camada *dropout* com $p = 0,5$. Os vértices em cor cinza foram desativados na iteração e os outros continuaram ativos.

2.9 Comitê de Classificadores

O Comitê de Classificadores é uma abordagem com a ideia de utilizar diversos classificadores para realizar uma predição em conjunto. A Figura 17 representa um exemplo de um comitê.

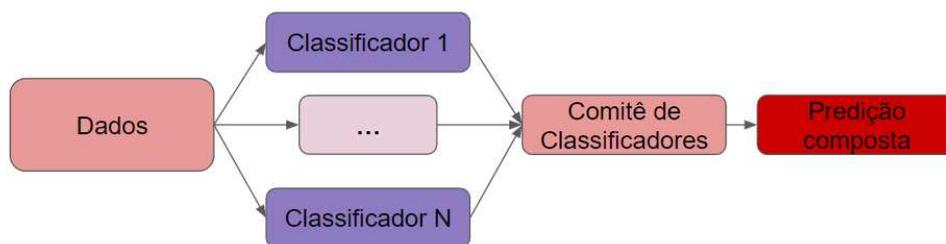


Figura 17 – Exemplo de um Comitê com N classificadores.

Para utilizar desse método, algumas características precisam ser pré-definidas para cada classificador (GENTLE; HÄRDLE; MORI, 2012):

- **Tipos de algoritmos:** Os classificadores podem se diferenciar em relação aos seus algoritmos de classificação. Por exemplo, para um comitê com 3 classificadores, uma distribuição válida seria definir o primeiro como uma Árvore de Decisão, o segundo uma Floresta Aleatória e o último uma Rede neural.

- **Tipos de características:** Os classificadores não necessariamente usam o mesmo conjunto de características para classificação, pode-se usar um conjunto específico para cada classificador do comitê.
- **Amostras utilizadas:** A quantidade de amostras utilizadas pode ser diferente para cada classificador.

Além disso, é preciso definir um combinador, este decidirá quando cada classificador será utilizado e como os resultados serão combinados ao final. Existem diversas maneiras diferentes para combinar os classificadores, a depender de como a arquitetura foi modelada:

- **Arquitetura paralela:** Nesta arquitetura, cada classificador desempenha de maneira independente e todos os resultados são combinados ao final.
- **Arquitetura cascata:** Nesse modelo, os classificadores são executados de maneira sequencial, os resultados são passados ao próximo modelo em fila e as possibilidades de classes vão diminuindo no decorrer dos classificadores.
- **Arquitetura hierárquica:** Nessa abordagem, os classificadores são agrupados por hierarquia com condicionais, como se fosse uma árvore de decisão, e todos os nós são os classificadores, mas nem todos necessariamente serão usados no processo em razão da presença das condicionais.
- **Arquitetura híbrida:** Nesse método, pode-se misturar as diferentes arquiteturas definidas anteriormente para a construção de um modelo híbrido. Por exemplo, é possível que uma parte do modelo utilize arquitetura hierárquica, ou seja, uma árvore com vários classificadores e o restante do modelo seja com classificadores agrupados sequencialmente, como na arquitetura em cascata.

No que se refere à aplicação do *ensemble learning*, algumas técnicas são conhecidas na literatura, como por exemplo o *Bagging* e o *Boosting* (GENTLE; HÄRDLE; MORI, 2012).

No *Bagging*, vários classificadores são selecionados para compor o comitê e cada um desses classificadores é treinado em paralelo com um subconjunto dos dados independentes. Assim, para fazer a classificação de um novo dado, é feita uma predição sobre ele em cada um dos classificadores e, por fim, é realizada uma agregação sobre as saídas de todos os classificadores por meio de moda ou média. O *Random Forest* é um exemplo de *ensemble model* que utiliza a abordagem *Bagging*, ele faz uso de várias árvores de decisão em paralelo e depois realiza uma agregação das saídas das árvores.

No *Boosting*, vários classificadores são colocados em cascata. O primeiro classificador recebe o conjunto de dados atribuindo peso igual para todos os pontos do conjunto.

Após realizar o treinamento, é atribuído peso maior aos dados em que o classificador errou e então esse conjunto com os novos pesos é passado para o próximo classificador. Esse processo se repete até o último classificador da cascata, após isso tem-se um modelo treinado resultante da combinação de todos os classificadores anteriores.

2.10 Métricas de avaliação

Após a obtenção de resultados, é necessário fazer uso de métricas de avaliação para medir o quanto o método se sobressaiu sobre o problema.

A saída de diversos algoritmos de classificação produz uma matriz de confusão que permite analisar algumas métricas de validação, como, por exemplo, a acurácia (Equação 2.19), a especificidade (Equação 2.20), a precisão (Equação 2.21), a sensibilidade (Equação 2.22) e o *f1-score* (Equação 2.23).

$$\text{acuracia } (ac) = \frac{VP + VN}{VP + VN + FP + FN}. \quad (2.19)$$

$$\text{especificidade } (es) = \frac{VN}{VN + FP}. \quad (2.20)$$

$$\text{precisao } (p) = \frac{VP}{VP + FP}. \quad (2.21)$$

$$\text{sensibilidade } (se) = \frac{VP}{VP + FN}. \quad (2.22)$$

$$\text{f1 - score } (f1) = \frac{2 \times VP}{2 \times VP + FN + FP}. \quad (2.23)$$

A matriz de confusão é uma matriz com valores quantitativos que nos ajudam na classificação de amostras, são eles: Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN). Refletindo esses valores nas amostras da displasia, temos que:

- **VP:** É a quantidade de pacientes infectados com displasia.
- **VN:** é a quantidade de pacientes que não estão com displasia detectados como não infectados com a displasia.
- **FP:** é a quantidade de pacientes que não estão com displasia, mas que foram, erroneamente, detectados com displasia.
- **FN:** é a quantidade de pacientes com displasia, mas que foram, erroneamente, detectados como não infectados com displasia.

2.11 Trabalhos relacionados

As displasias, quando estimuladas fomentadas por estímulos como tabaco e álcool, por exemplo, podem desenvolver um potencial maligno, crescendo consideravelmente as chances de evolução para câncer oral de 65%. Desse modo, é evidente a necessidade de investigação de técnicas computacionais para auxiliar o especialista com diagnósticos mais precisos, uma vez que a taxa de mortalidade dessas lesões é alta.

O estudo [Baik et al. \(2014\)](#) apresentou um método para identificar lesões com alto risco de evolução para câncer. Nesse método, foram utilizados algoritmos baseados na Floresta Aleatória para fazer a segmentação das imagens. Após a etapa de segmentação foi utilizado outro algoritmo baseado em RF para realizar a classificação binária entre tecido saudável e tecido não saudável, além disso, utilizou-se da abordagem de validação cruzada *k-fold*, empregando $k = 10$. O banco de imagens foi composto de 133 regiões com 29 tecidos saudáveis e 104 tecidos não saudáveis. O método obteve uma acurácia de 80% na classificação dos tecidos.

O estudo em [Adel et al. \(2018\)](#) desenvolveu um sistema CAD para classificar imagens histológicas de displasia epitelial oral. Para esse trabalho, utilizou-se um banco de imagens com 39 imagens de tecido representando displasia e 7 imagens de tecidos saudáveis. Para extrair características, foram usados os algoritmos: *Scale Invariant Feature Transform*, *Speeded Up Robust Features* e ORB (do inglês, *Oriented FAST and Rotated BRIEF*). Isso possibilitou a extração de 16 características. Para finalizar, foram utilizados os algoritmos *SVM* e *KNN* para realizar a classificação binária das imagens. O algoritmo *SVM* obteve o maior resultado, atingindo uma acurácia de 92,8% quando utilizado com as características ORB. Utilizando o algoritmo *KNN*, foi obtido uma maior acurácia de 78,6% usando as características ORB.

No estudo [Huttunen et al. \(2020\)](#), um método foi proposto para detecção de displasia em tecidos epiteliais juntamente a lesões na pele humana. Esse trabalho utilizou um banco de imagens com 358 imagens histológicas para avaliar o método. Para fazer a classificação, foi utilizado a técnica de *transfer-learning* utilizando uma rede neural *GoogLeNet*. Na classificação, foram aplicadas apenas duas classes: saudável e tecido com displasia. A classificação teve uma acurácia de 94.2%. Esse estudo apresentou um método que permite a identificação de displasia, entretanto, o método não fazia a distinção de diferentes níveis de displasia.

Por fim, no trabalho [Silva et al. \(2022a\)](#), foi desenvolvido um sistema CAD para classificação de displasia epitelial oral. Após isso, algumas operações morfológicas foram realizadas para remover falsos positivos e áreas negativas. Em seguida, 23 características morfológicas e não morfológicas foram extraídas, tais como: área, orientação, entropia e solidez. Para classificar, foi utilizado um classificador polinomial para distinguir as imagens

de acordo com seu nível de lesão. Essa abordagem foi utilizada juntamente a um banco com 296 imagens de língua de camundongo, onde 9155 celular nucleicas foram identificadas e analisadas. Para avaliar os métodos, foram utilizadas métricas como acurácia e área sobre a curva ROC. O estudo atingiu acurácias que variam de 88,92% a 90,35% e a etapa de classificação atingiu áreas ROC variando entre 0,88 a 0,97.

Após analisar os resultados citados, é notório que a literatura sobre displasia oral está em constante crescimento. Entretanto, há ainda vários desafios a serem explorados. Diante disso, estudos baseados em *deep learning* sobre as displasias podem trazer resultados promissores para a literatura.

3 Desenvolvimento

Este capítulo apresenta detalhes do banco de imagens utilizado para avaliações quantitativas da classificação proposta. Também são apresentados os algoritmos empregados para cálculo dos descritores da geometria fractal, representação 2D e classificação dos graus das lesões presente na cavidade oral.

3.1 Visão geral

Este trabalho apresenta uma ferramenta computacional para classificação de lesões histológicas da cavidade oral baseado em descritores fractais e comitê de classificadores com abordagens de CNNs em cascata. As principais etapas do estudo são: i) emprego da geometria fractal para extração de descritores de textura e representação em 2D das informações; ii) utilização das imagens construídas a partir dos descritores fractais nos *backbones* de CNNs para extração de características e classificação; iii) estudo de modelo de comitê de classificadores para avaliação do desempenho da abordagem. Na Figura 18 são apresentadas as etapas utilizadas para a classificação das imagens histológicas da cavidade oral.

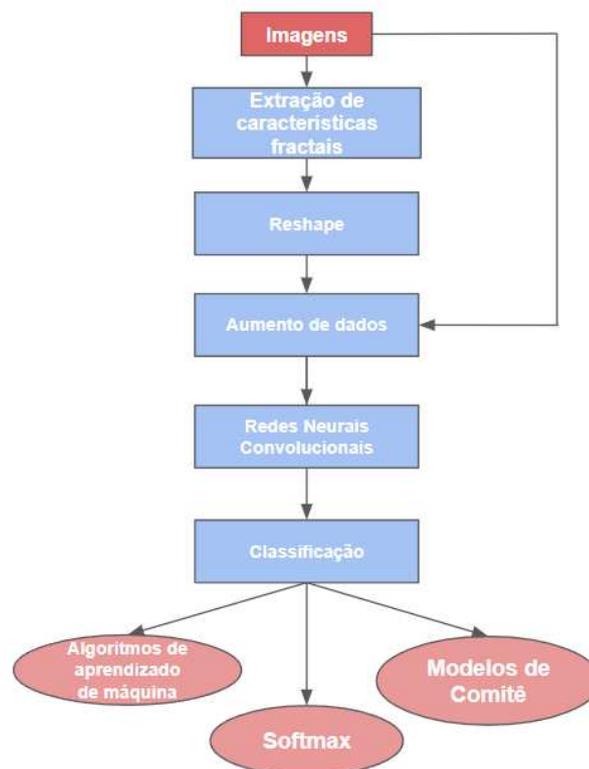


Figura 18 – Sequência metodológica utilizada para a classificação de lesões da cavidade oral.

Os métodos apresentados na Figura 18 foram implementados nas linguagens Python e Matlab. Os experimentos foram realizados em um *computador* com processamento de 3,40 GHz (Intel Core i5-4670k) e 12GB de memória RAM. O código-fonte com as informações a respeito das implementações estão disponíveis na URL:

<https://github.com/tgoMota/TCC-Thiago-Mota-Carvalho/>.

3.2 Banco de Imagens

O banco de imagens é composto de imagens microscópicas de lesões da língua de camundongos obtidas por meio de experimentos aprovados pelo Comitê de Ética da Universidade Federal de Uberlândia com o número de protocolo 038/09. Essas imagens também foram utilizadas nos trabalhos apresentados por (SANT'ANA et al., 2011), (FARIA et al., 2011), (SILVA et al., 2019), (SILVA et al., 2022a) e (SILVA et al., 2023).

As lâminas foram obtidas com o uso do protocolo de indução de lesões em língua de camundongos. Para esse experimento foram utilizados 30 animais com peso variando de 20 a 23 gramas e com idade entre 6 e 8 semanas. Todos os animais foram distribuídos de maneira aleatória em dois subgrupos. Os camundongos foram abrigados em gaiolas com material de plástico a uma temperatura de 22°C, com ciclo claro/escuro de 12h na Universidade Federal de Uberlândia, no Centro de Bioterismo. O experimento adotou os critérios propostos pelos autores em (TANG et al., 2004). Antes dos experimentos, o carcinógeno 4NQQ foi diluído em propilenoglicol na concentração de 5 mg/ml e, na sequência, diluído em água filtrada até alcançar uma concentração final de 100 mg/ml. Após todos os animais serem sacrificados, suas línguas foram removidas e colocadas em solução de formalina tamponada a 4% por 24 horas. Posteriormente, foi conduzido um processamento das línguas e foram incluídas em parafina, recortadas e depois coradas por Hematoxilina e Eosina (H&E) para estudo histopatológico das lâminas. Ao final, um patologista realizou a classificação das lâminas em tecido saudável, displasia leve, displasia moderada e displasia grave.

Com uso de um microscópio óptico Leica DM500, utilizando magnificação 400×, as imagens histológicas digitais foram obtidas e armazenadas em arquivos com modelo de cores RGB e formato .TIFF, quantificação de 8 bytes e com resolução de 2048 x 1536 pixels. Num grupo de 43 lâminas as imagens foram digitalizadas e foram recortadas as regiões de interesse (do inglês, *Region of Interest* - ROI), sendo 148 ROIs para cada classe de tecido com tamanho 450 × 250 *pixels* e totalizando 592 ROIs. Na Figura 19 é apresentado exemplos de ROIs digitalizadas obtidas de cada uma das classes investigadas no trabalho.

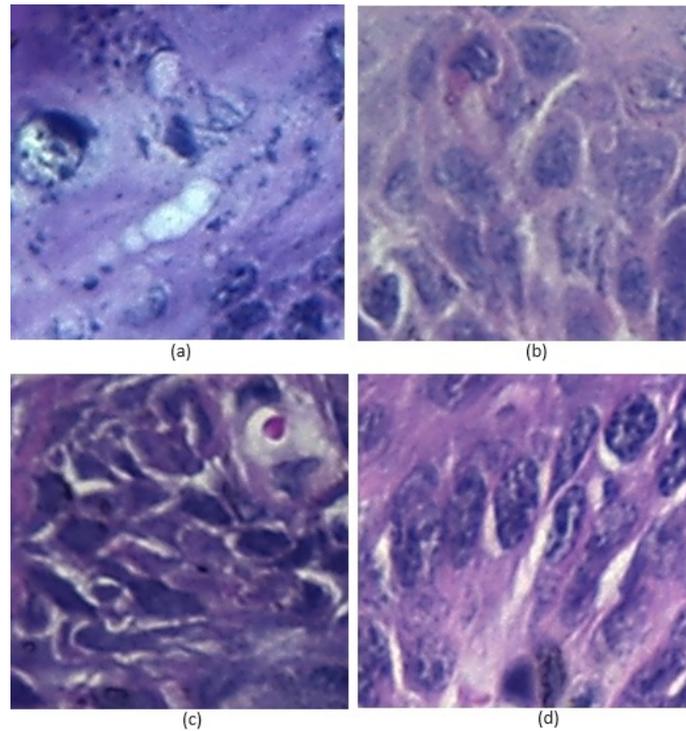


Figura 19 – Exemplos de tecidos histológicos da cavidade bucal digitalizados: (a) saudável, (b) displasia leve, (c) displasia moderada e (d) displasia severa. **Fonte :** (SILVA et al., 2019)

3.3 Representação 2D de Atributos Fractais

3.3.1 Geometria Fractal

Para obtenção dos descritores baseados em geometria fractal foram aplicadas as matrizes de probabilidades, essas matrizes são responsáveis por conter as probabilidades da imagem de uma região quadrada preenchida. O processo de análise do preenchimento foi realizado por meio da relação de distância entre valores de pixels e o tamanho da área conforme descrito em (CANTRELL, 2000). Essas matrizes de probabilidades foram obtidas com o método *gliding-box*, como descrito na Seção 2.5, que consistiu em deslocar uma janela quadrada por toda a imagem e analisar os pixels correspondentes dentro ou fora da caixa (ALLAIN; CLOITRE, 1991).

Utilizando o algoritmo *gliding-box*, as imagens foram divididas em diferentes escalas criando um conjunto de matrizes para cada região da imagem em cada valor de L que varia entre 3 e 41 com um incremental de 2. Esse método foi empregado usando dados obtidos do estudo de Ribeiro et al. (2019). O valor 41 escolhido como L_{max} foi com objetivo de não gerar distorções no redimensionamento das imagens antes de prepará-las para dar como entrada para as redes neurais. O incremento no valor de unidade 2 em cada iteração permitiu garantir quantificações em 20 escalas diferentes. A partir destes procedimentos os descritores de dimensão fractal, lacunaridade e percolação, ambos

multiescalas e multidimensionais, foram calculados por meio da estratégia de matriz de probabilidades (IVANOVICI; RICHARD, 2010).

Com a obtenção da matriz de probabilidade foram obtidas 20 características por função (D(L), $\Lambda(L)$, C(L), Q(L) e M(L)), resultando em cerca de 100 descritores. Para finalizar, foi utilizada a abordagem da dimensão fractal descrita na Seção 2.5 para cálculo de diferentes distâncias Δ entre pixels da região.

Após a aplicação do método da dimensão fractal, tem-se um conjunto de matrizes com valores binários (0 ou 1), nelas o valor 1 representa os pixels que satisfizeram o critério de distância (Δ).

Por fim, cerca de 300 características locais foram extraídas e utilizadas na etapa de representação em que o método de gráfico de recorrência, descrito na Seção 3.3.2, foi empregado. Além das características locais, se pode obter as características globais da geometria fractal com esses descritores, contudo, elas não foram utilizadas neste trabalho.

3.3.2 Representação com Método de Gráfico de Recorrência

Na seção 3.3.1 foram descritos os procedimentos para extração de características baseadas em geometria fractal com os descritores de dimensão fractal, lacunaridade e percolação, ambos multiescalas e multidimensionais. Essas informações numéricas em 1D não podem ser dadas como entrada para uma *CNN* com camadas convolucionais avaliadas em operações 2D. Nesse caso, são necessários novos procedimentos para uma representação dos dados em um plano 2D. Neste estudo, para que essa operação fosse realizada, foi empregado o método de gráfico de recorrência (do inglês, *recurrence plot* - RP).

A técnica de RP proposta em Eckmann et al. (1995) consiste na obtenção de projeção de eventos repetidos em 2D ou 3D. Essa estratégia é costumeiramente aplicada em diversos contextos como, por exemplo, a detecção da doença de Parkinson como descrito em (AFONSO et al., 2019). A aplicação dessa técnica em um vetor contendo N características, resulta em uma matriz quadrada $N \times N$. Para obtenção de cada elemento da matriz emprega-se a Equação 3.1:

$$L_{i,j} = \| x_i - x_j \| \quad \forall i,j \in 1..N . \quad (3.1)$$

onde x_i e x_j representam, respectivamente, a i 'ésima e a j 'ésima característica no vetor em análise. Essa abordagem é baseada no trabalho de Roberto et al. (2021) em que esse processo pode ser observado na Figura 20.

3.4 Rede Neural Convolutacional

Com base na abordagem apresentada pelos autores em Kathamuthu et al. (2023), modelos de redes pré-treinadas são usadas como entrada para um segundo modelo de rede

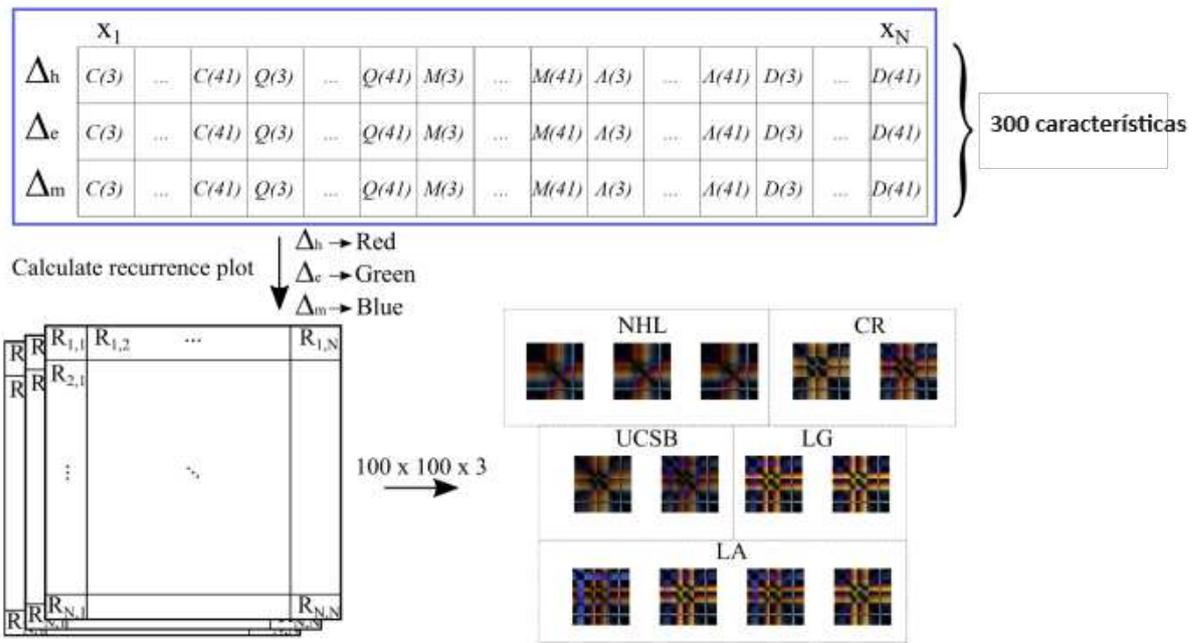


Figura 20 – Método de gráfico de recorrência sendo aplicado ao vetor de características para gerar as imagens. Fonte: (ROBERTO et al., 2021)

customizado. Neste estudo, baseado na proposta de Kathamuthu et al. (2023), as saídas das camadas convolucionais de modelos pré-treinados foram usados como entrada para um novo *backbone* de CNN para classificação multi classes.

Para esta pesquisa, inicialmente, foram empregadas as redes *ResNet50* (CELANO, 2021) e a *MobileNet* (HOWARD et al., 2017) com abordagem de transferência de aprendizado. O método de transferência de aprendizado consiste em aproveitar os pesos de um modelo que foi treinado em um outro banco de dados. Neste trabalho, os pesos dos modelos pré-treinados foram referentes a um treinamento usando o banco de imagens *ImageNet* (DENG et al., 2009). Em todos os testes com transferência de aprendizado, o modelo pré-treinado foi usado como entrada para uma CNN customizada.

A *ResNet50* é um modelo de rede que se identifica como uma variação da *ResNet* tradicional, em que o modelo possui 48 camadas de convolução seguidas de camadas de *pooling*. Essa arquitetura é um modelo complexo e possui cerca de aproximadamente 25 milhões de parâmetros. Na Figura 21 é apresentada a *ResNet50* e sua estrutura de camadas. Nessa rede, os dados foram obtidos da camada *conv5-block3-out* e foram utilizados como entrada para o modelo customizado proposto neste trabalho.

Por outro lado, a *MobileNet* é um modelo de rede que foi projetada com o objetivo de oferecer uma rede neural que possa ser utilizada em equipamentos de baixo poder computacional, sendo eficiente e rápida. Esse modelo faz uso de camadas convolucionais do

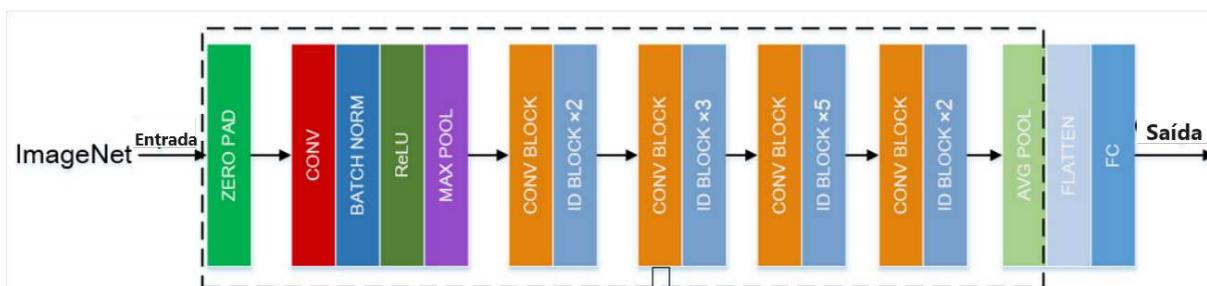


Figura 21 – Estrutura da arquitetura da *Resnet50* pré-treinada com *ImageNet* **Fonte:** (TANG; LI; WANG, 2020)

tipo *depthwise separable convolution* (HOWARD et al., 2017) e cada uma dessas camadas convolucionais é seguida por uma camada de normalização *batch* e uma função de ativação *ReLU*. Algumas das menores versões da *MobileNet* possuem menos que 2 milhões de parâmetros. Na Figura 22 é mostrada a estrutura da *MobileNet*. Nesta análise, os dados foram retirados da camada *conv-pw-13-relu*, em que o mapa foi empregado sobre o novo modelo customizado.

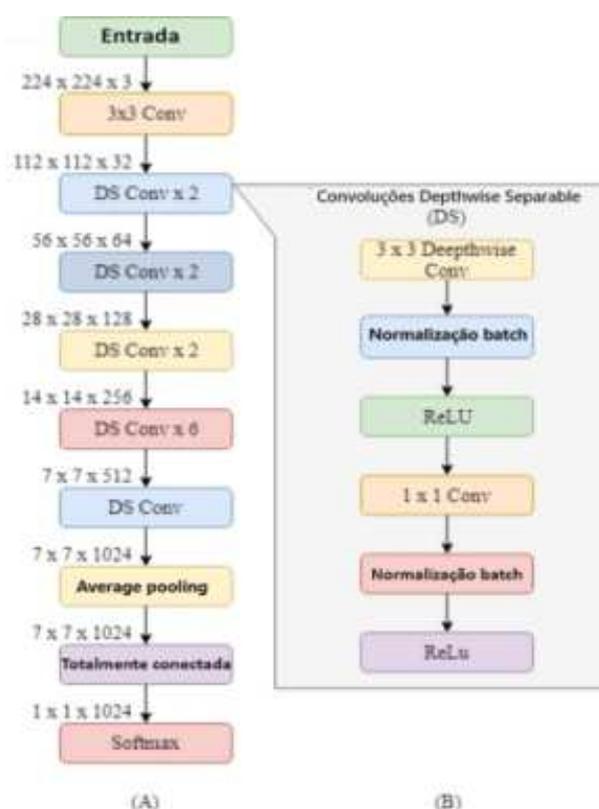


Figura 22 – Estrutura da arquitetura da *MobileNet* **Fonte:** (PHIPHITPHATPHAISIT; SURINTA, 2020)

Por fim, o novo modelo em cascata foi utilizado, esse, constituído de 2 camadas convolucionais, seguidas de camadas de normalização *batch* e camadas de *pooling* com

valor máximo. Após todas as camadas de normalização *batch*, foi utilizada uma função de ativação *ReLU* para eliminar valores negativos (AGARAP, 2018). No final do modelo foi utilizada uma camada *flattening* seguido de uma camada totalmente conectada com um regularizador *dropout* com uma taxa de 0.02. Esse modelo foi compilado utilizando o otimizador *adam*, baseado em (KINGMA; BA, 2014), com uma taxa de treinamento de 0.001, e utilizando uma função de ativação *softmax* de *categorical crossentropy* (ZHANG; SABUNCU, 2018). A Tabela 1 mostra as camadas do modelo proposto.

Tabela 1 – Estrutura do modelo proposto customizado com as camadas e saída das operações em cada camada.

Camada	Saída
convolucional 2D	(4, 4, 32)
normalização <i>batch</i>	(4, 4, 32)
função de ativação <i>ReLU</i>	(4, 4, 32)
polimento máximo	(2, 2, 32)
convolucional 2D	(2, 2, 64)
normalização <i>batch</i>	(2, 2, 64)
função de ativação <i>ReLU</i>	(2, 2, 64)
polimento máximo	(1, 1, 64)
<i>flatten</i>	(64)
totalmente conectada	(128)
normalização <i>batch</i>	(128)
função de ativação <i>ReLU</i>	(128)
regularizador <i>dropout</i>	(128)
totalmente conectada	(2)
função de ativação <i>softmax</i>	(2)

Com o modelo proposto foi realizado um treinamento por 100 épocas com o método de amostragem *holdout*. O método *holdout* é simples, consiste em dividir o conjunto de dados em treino e teste. Para esse trabalho, além dos dados de treino e teste, também foram separados os dados de validação. A separação dos dados ocorreu na forma 80/10/10, em que 80% das imagens representam o conjunto de treino, 10% representam o conjunto de teste e os outros 10% representam o conjunto de validação. Essa divisão foi escolhida pelo motivo do banco de imagens utilizado não ser muito grande, portanto, a quantidade de dados separados para treinamento deve ser grande o suficiente para construir um modelo de classificação robusto sem a presença de *overfitting*.

Após o treinamento, para cada imagem foram extraídas cerca de 64 características resultantes da saída da camada *flatten* do novo modelo. Essas características foram

armazenadas em um vetor para serem usadas na etapa de classificação com algoritmo de aprendizado de máquina.

3.5 Operações de Aumento de Dados

As imagens do banco foram redimensionadas para 128×128 *pixels* mantendo o padrão do modelo de cores RGB. Foi escolhido a dimensão 128×128 *pixels*, pois é um tamanho em que não há muitas perdas de informações, e devido a capacidade de processamento do equipamento usado para os experimentos. Após o redimensionamento, as imagens foram separadas e rotuladas em quatro classes: Saudável, Displasia Leve, Displasia Moderada e Displasia Grave. Na Tabela 2, tem-se um resumo do conjunto de dados.

Tabela 2 – Conjunto de dados investigados nesse trabalho para classificação das lesões.

Classe	Quantidade	Formato
Saudável	148	(128,128,3)
Displasia Leve	148	(128,128,3)
Displasia Moderada	148	(128,128,3)
Displasia Grave	148	(128,128,3)

Como o número de imagens originais por classe é muito pequeno (148 amostras para cada classe), o processo de treinamento do modelo torna-se complicado para a aprendizagem e generalização. Como é inviável conseguir mais imagens originais provenientes do mesmo conjunto de lâminas, uma das alternativas é aplicar técnicas de aumento de dados, como descrita na Seção 2.7. No aumento de dados foram aplicadas as seguintes operações: rotação em 90 graus, espelhamento horizontal e espelhamento vertical. Então, cada imagem no banco de dados gerou 3 novas imagens. Na Tabela 3 tem-se uma sumariação do novo conjunto de imagens após aplicação das técnicas de aumento de dados. Essas novas imagens foram dadas como entrada para as CNNs para os procedimentos de extração de características e classificação.

Tabela 3 – Número de imagens após aplicação das operações de aumento de dados.

Classe	Quantidade	Formato
Saudável	592	(128,128,3)
Displasia Leve	592	(128,128,3)
Displasia Moderada	592	(128,128,3)
Displasia Grave	592	(128,128,3)

3.6 Comitê de Classificadores

Para o comitê de classificadores, optou-se pela abordagem *bagging*, visto na Seção 2.9, em que vários classificadores são utilizados em cascata para tratar um problema de classificação. A abordagem foi selecionada devido ao desempenho observado nos estudos AlJame et al. (2020) e Das et al. (2021). Para essa pesquisa, foram criados dois modelos de comitês de classificação, em que o primeiro comitê foi composto por três CNNs: *MobileNet*, *ResNet50* e *ResNet50V2*. Embora a *ResNet50* e a *ResNet50V2* sejam modelos iguais, porém, de versões diferentes, a *ResNet50V2* pode apresentar valores melhores em certos casos, como apresentado no estudo Hieu e Hien (2020) e, por isso, foi escolhido para compor o primeiro comitê.

Para o segundo comitê, os algoritmos de aprendizado de máquina (do inglês, *machine learning* - ML) descritos na Seções 2.6, 2.6.3 e 2.6.4 foram empregados baseado nas abordagens: Árvore de Decisão (do inglês, *decision tree* - DT), Floresta Aleatória (do inglês, *random forest* - RF) e Regressão Logística (do inglês, *logistic regression* - LR). Todas as etapas dos algoritmos foram descritas na Seção 2.6.

Para os algoritmos ML, no caso, da RF foram utilizados os valores padrões da biblioteca, são eles: o número de árvores na floresta é igual a 100, profundidade máxima das árvores é igual a *None* (as árvores se expandem até atingir vértices folhas) e máximo número de vértices folhas é *None* (não há limite para a quantidade de vértices folhas). Para a DT, também foram utilizados os valores padrões da biblioteca, são eles: profundidade máxima da árvore é *None* (expande sem limite de profundidade), máxima quantidade de vértices folhas é *None*, entre diversos outros atributos também definidos com o valor padrão da biblioteca. Para a LR, os valores padrões da biblioteca também foram utilizados, como: número máximo de iterações até convergir sendo 100, pesos das classes sendo 1, entre outros.

Para os comitês propostos, torna-se necessário definir a operação estatística que será feita na junção das predições dos classificadores. Duas abordagens que podem ser empregadas são: a moda, em que o rótulo que possuir mais votos pelos classificadores será o rótulo de saída do algoritmo; ou a média, em que é feita uma média das probabilidades atribuídas a cada rótulo por todos os classificadores, o rótulo com maior probabilidade final será o resultado da classificação. Neste estudo, as duas abordagens foram utilizadas e avaliadas. Na Figura 23 é representado como os comitês foram construídos para o desenvolvimento deste trabalho.

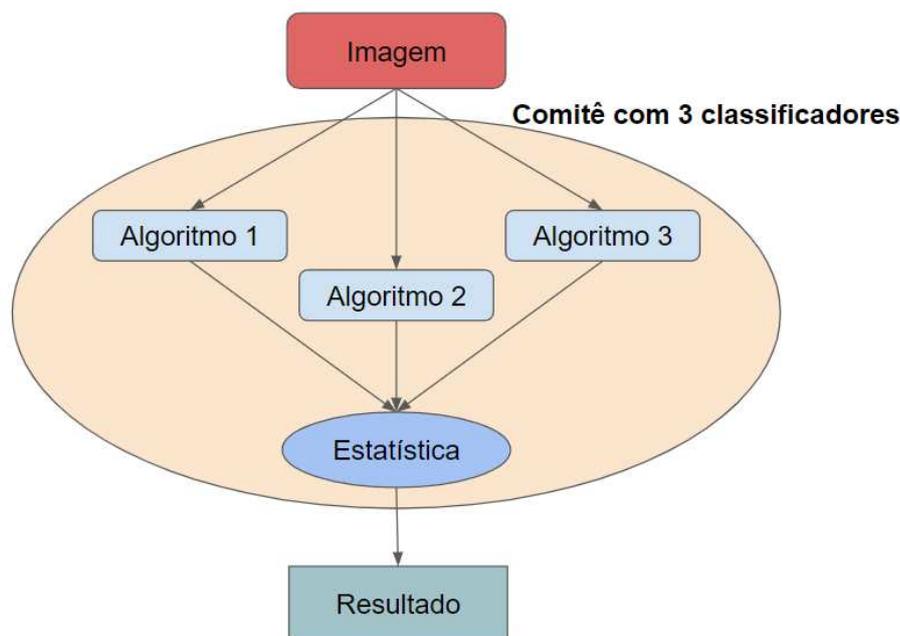


Figura 23 – O comitê formado por 3 algoritmos (Algoritmo 1, 2, 3), o qual recebe o resultado de cada modelo e realiza a operação estatística (moda ou média).

3.7 Avaliação de Desempenho das Abordagens

Para a execução dos experimentos, o banco de imagens foi avaliado na abordagem multi classes entre os estágios do grau da displasia e tecidos saudáveis. Na investigação foram avaliados os algoritmos de CNNs e ML de forma separada, denominados de *baseline*. Esses modelos foram explorados com objetivo de avaliar o comportamento dos dados sem uso de comitês. As representações e imagens digitais foram aplicadas nas redes neurais convolucionais (CNNs) para classificação e obtenção de características do *flatten* antes das camadas totalmente conectadas no modelo proposto. Os comitês investigados exploraram os dados sobre a abordagem de modelos de algoritmos de ML e também sobre a composição de CNNs com o propósito de investigar o comportamento das imagens digitais e representações sobre essas aproximações. Neste estudo, os comitês foram denominados Comitê-CNN, para o modelo composto das redes *MobileNet*, *ResNet50* e *ResNet50V2*, e Comitê-ML, para a composição dos algoritmos DT, RF e LR.

4 Resultados

Nessa seção são apresentados os resultados obtidos dos modelos de redes neurais convolucionais (do inglês, *convolucional neural network* - CNN) com a abordagem em cascata, assim como, o desempenho dos comitês de classificação investigados para os algoritmos de aprendizado de máquina (do inglês, *machine learning* - ML) e as CNNs.

4.1 Avaliação de Desempenho dos Algoritmos de ML para Lesões Histológicas

Neste estudo, inicialmente, os algoritmos de ML foram avaliados sobre os mapas de características obtidas das camadas finais das CNNs customizadas conforme descritos na Seção 3.4 do trabalho. A Tabela 4 apresenta a classificação com os algoritmos ML aplicados em Imagens Digitais (ID) de displasia e representações 2D obtidas com a técnica Gráfico de Recorrência (GR) extraídas dos descritores fractais.

O algoritmo floresta aleatória (do inglês, *random forest* - RF) obteve um desempenho superior aos demais modelos quando empregado nas imagens de GR para a classificação com a amostragem *holdout*. Essa abordagem obteve uma acurácia de 100% e *F1-score* de 1,00 para os dados utilizados para o teste, indicando que esse modelo se comportou positivamente no que se refere ao conjunto de imagens utilizadas. No entanto, com o uso das imagens da representação GR esses valores foram inferiores. Nota-se ainda que no caso da métrica *F1-score*, a qual analisa o desbalanceamento entre as classes de lesões investigadas, o valor foi o menor em relação aos demais experimentos. Diferente da métrica acurácia, a *F1-score* mostrou que o desbalanceamento entre as classes das imagens ID foi detectado nos experimentos na fase de teste. Dentre os classificadores ML, a árvore de decisão (do inglês, *decision tree* - DT) quando aplicada às IDs, foi obtido uma acurácia de 85,89%, com uma diferença de quase 15% quando comparado a RF com as imagens de GR. Esse algoritmo obteve um valor de 0,85 para o *F1-score*, o que significa que os

Tabela 4 – Desempenho dos Algoritmos de ML para Classificação das Lesões Histológicas

Algoritmos de ML	Imagem	Acurácia (%)	F1-Score
DT	ID	85,89	0,85
	GR	98,75	0,98
RF	ID	89,40	0,78
	GR	100,00	1,00
LR	ID	86,75	0,86
	GR	99,15	0,99

Tabela 5 – Avaliação das imagens com uso dos modelos de CNNs em cascata.

Classificador	Método	Acurácia (%)	F1-Score
MobileNet	ID	96,61	0,78
	GR	100,00	1,00
ResNet50	ID	95,76	0,95
	GR	99,15	0,99
ResNet50V2	ID	94,49	0,94
	GR	99,15	0,99

valores de acurácia e sensibilidade foram próximos, sem divergências. Os resultados alcançados com o algoritmo regressão logística (do inglês, *logistic regression* - LR) ficaram no intervalo entre os métodos anteriores.

4.2 Resultados com os Modelos CNNs em Cascata

Na Tabela 5 são apresentados os resultados da classificação com os modelos CNNs em cascata com os dados atingidos da ID e do GR. Com a rede *MobileNet* na entrada do modelo customizado a aproximação proposta obteve, utilizando o método GR, uma acurácia de 100,00% e um *F1-score* de 1,00. Isso indica que o classificador aprendeu a generalizar os dados conseguindo uma classificação relevante sem falsos positivos e falsos negativos. Com base nesses resultados, na Figura 24 são mostradas as curvas de acurácia durante o treinamento e validação com o modelo *MobileNet* para as imagens obtidas pelo método GR, em que, na cor azul, os gráficos mostram a acurácia com as imagens separadas para treinamento e, utilizando a cor verde, a acurácia com dados de validação. Analisando esse gráfico destaca-se que o modelo obteve variações nesse processo de aprendizagem até a época 30 e a acurácia começou a estabilizar a partir da época 40.

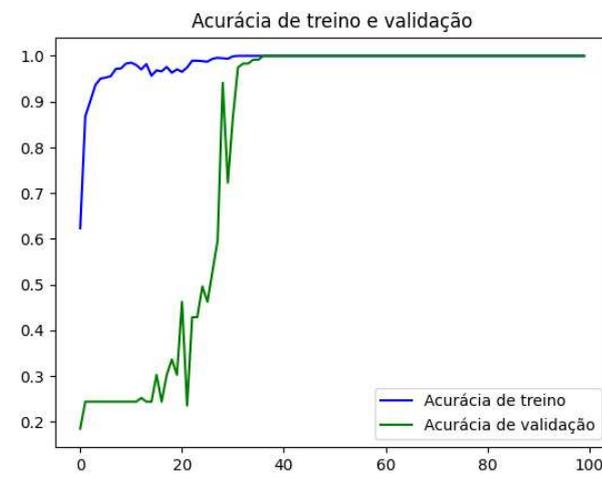


Figura 24 – Treinamento em 100 épocas utilizando a rede *MobileNet* com as imagens obtidas pelo método GR.

No caso das imagens ID, os melhores resultados ocorreram com a *ResNet50*, em que a métrica *F1-score* foi de 0,95. A *ResNet50V2* foi o modelo que obteve desempenho muito próximo ao modelo da Versão 1 da *ResNet50*. Ambos os modelos têm os blocos residuais, onde a saída de um bloco é a soma da entrada original com a saída do bloco. No entanto, a *ResNetV2* tem algumas modificações para melhorar a eficiência e o desempenho, em que a operação de *Batch Normalization* antes das ativações permite um número menor de parâmetros. A Figura 25 mostra as curvas de acurácia na fase de treino dos modelos *ResNet50* e *ResNet50V2* utilizando as imagens ID. Embora produziram acurácias similares nos dados de teste, a segunda versão da *ResNet50* não apresentou curvas de acurácia tão boas, não houve linearidade ao longo das épocas e vários mínimos locais, o que pode indicar a presença de *overfitting*. O modelo *ResNet50* apresentou curvas que demonstram um rápido aprendizado, pois, por volta da época 25, o modelo estabilizou em uma acurácia oscilando entre 95,00 e 96,00. Embora os modelos tenham a mesma acurácia geral, a *ResNet50* na versão 1 demonstrou que conseguiu generalizar os dados com menor dificuldade.

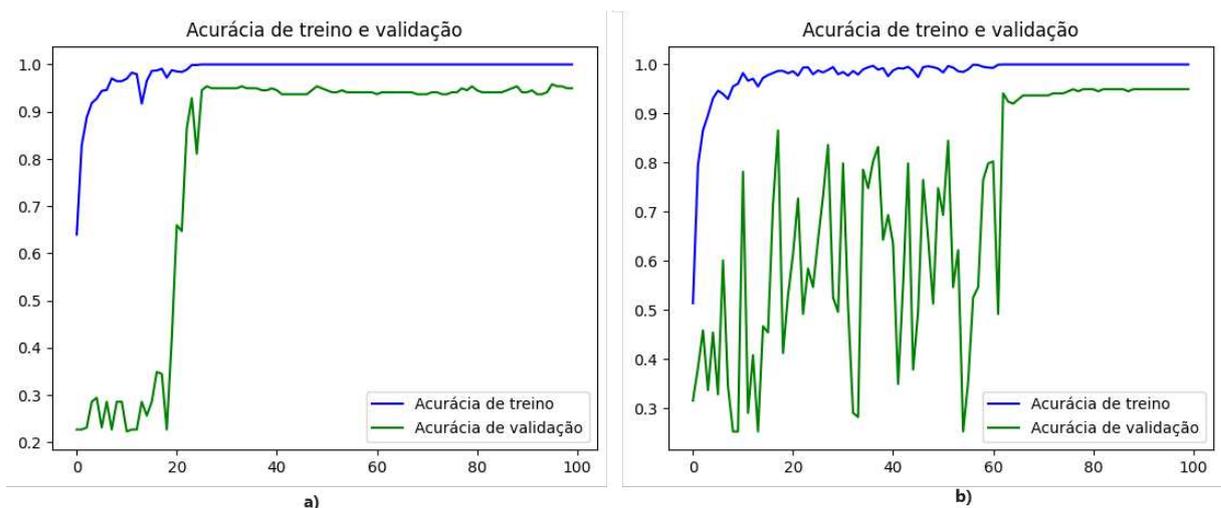


Figura 25 – Gráficos que mostram o processo de treinamento da (a) *ResNet50* e (b) *ResNet50V2* utilizando as imagens obtidas pelo método ID.

4.3 Investigação dos Comitês de Classificadores

Na Tabela 6 são apresentados os resultados com os valores obtidos na investigação dos comitês utilizando os algoritmos ML e CNN. É possível notar que, com o uso de comitês, todos os experimentos utilizando o método GR obtiveram uma taxa de acurácia de 100%. Por outro lado, os experimentos utilizando os dados ID resultaram em valores inferiores, que variaram de 94,49% a 96,18%. O valor mais alto utilizando o método ID com os comitês ocorreu no modelo Comitê-CNN com acurácia de 96,18%, mas quando

Tabela 6 – Desempenho dos modelos baseados em comitê de classificadores para as abordagens ML e CNN.

Comitê	Operação	Método	Acurácia (%)	F1-Score
Comitê-CNN	Média	ID	94,49	0,92
		GR	100,00	1,00
	Moda	ID	96,18	0,96
		GR	100,00	1,00
Comitê-ML	Média	ID	95,53	0,94
		GR	100,00	1,00
	Moda	ID	94,49	0,94
		GR	100,00	1,00

comparado com a Tabela 5 a rede *MobileNet* foi inferior para o mesmo modelo de imagens. Nota-se ainda que o desempenho entre os dois modelos de comitê resultaram em valores próximos para as duas abordagens investigadas.

4.4 Análise de Desempenho com Estudos da Literatura

A Tabela 7 apresenta uma visão geral dos resultados obtidos em relação a outros métodos que empregam técnicas de visão computacional e inteligência artificial para as imagens histológicas da cavidade oral. Nessa tabela, todos os estudos foram feitos sobre a displasia oral, mas utilizaram diferentes métodos, tais como, *Oriented FAST and Rotated BRIEF* (ORB), Análise de Textura (AT), Características Morfológicas (CM), Características Aprendidas (CA), Gráfico de Recorrência (GR) e Imagens Digitais (ID) de displasia. Além disso, foram utilizados diferentes classificadores como, por exemplo, máquina de vetores de suporte (do inglês, *support vector machine* - SVM), *Hermite Orthogonal Polynomial* (HOP) e Comitê de Classificadores (CC).

Tabela 7 – Avaliação da metodologia proposta em relação às abordagens existentes para classificação de imagens orais presentes na literatura.

Estudo	Lesão	Método	Classificador	Acurácia
(ADEL et al., 2018)	Displasia Oral	ORB	SVM	92,60
(SILVA et al., 2022a)	Displasia Oral	AT e CM	HOP	92,40
(SILVA et al., 2022b)	Displasia Oral	CM	HOP	98,00
Método proposto	Displasia Oral	GR	CC	100,00
Método proposto	Displasia Oral	ID	<i>MobileNet</i>	96,61

5 Conclusão

Neste trabalho foi desenvolvido um sistema para detecção de displasias por meio da utilização de imagens histológicas aplicando técnicas de aprendizagem profunda e geometria fractal. Foram avaliados algoritmos *machine learning*, algoritmos *deep learning* e comitês de classificadores, fazendo uso de imagens histológicas de displasia e representações de descritores baseados em geometria fractal.

Os resultados obtidos mostraram que o uso de uma representação de imagem de características fractais, através do método de gráfico de recorrência em associação com os classificadores, contribuíram significativamente para a melhoria da precisão da classificação quando comparado a técnicas tradicionais que fazem apenas uso das imagens histológicas. Com o método proposto, foi possível obter uma acurácia de 100% sobre as imagens utilizadas no decorrer deste trabalho.

O uso de comitê de classificadores não se mostrou muito vantajoso para esta pesquisa. Quando empregado o comitê para os métodos com representações fractais, obteve-se uma acurácia de 100%. Nas abordagens apenas das redes em cascata esse valor foi de 100% de acurácia. Contudo, os testes sem uso de representação fractal, esperava-se uma melhoria nos valores das métricas para o uso do comitê. No entanto, os melhores valores para as métricas continuaram sendo as obtidas sem fazer uso dos comitês de classificadores.

Entre os desafios enfrentados, pode-se destacar o banco de imagens com pequeno número de amostras e baixo poder computacional para processamento na realização dos experimentos. Todavia, os resultados são promissores e fornecem uma contribuição para a literatura, uma vez que a representação das imagens com características fractais proporcionou aumento no desempenho da classificação. A utilização dos comitês de classificadores com as CNNs em modo cascata não obteve resultados muito relevantes para o domínio das lesões.

Em novos estudos outras abordagens podem ser exploradas com o objetivo de otimizar a metodologia apresentada. Com isso, em trabalhos futuros, será investigada a computação quântica juntamente às técnicas de inteligência artificial (GUPTA et al., 2017). De forma semelhante, se deve explorar outros bancos de imagens de displasia oral disponíveis pela comunidade científica.

Ressalta que, principalmente, as disciplinas de inteligência computacional, programação, lógica, inteligência artificial e mineração de dados foram essenciais na realização deste estudo. Elas contribuíram com um vasto conhecimento acerca da inteligência artificial e suas aplicações, além de proporcionar um aprofundado estudo em algoritmos complexos como: Árvore de Decisão, Floresta Aleatória e Redes Neurais Convolucionais.

Referências

- ADEL, D.; MOUNIR, J.; EL-SHAFFEY, M.; ELDIN, Y. A.; MASRY, N. E.; ABDELRAOUF, A.; ELHAMID, I. S. A. Oral epithelial dysplasia computer aided diagnostic approach. In: **2018 13th International Conference on Computer Engineering and Systems (ICCES)**. [s.n.], 2018. p. 313–318. Disponível em: <<https://doi.org/10.1109/ICCES.2018.8639452>>. Citado 2 vezes nas páginas 37 e 52.
- AFONSO, L. C.; ROSA, G. H.; PEREIRA, C. R.; WEBER, S. A.; HOOK, C.; ALBUQUERQUE, V. H. C.; PAPA, J. P. A recurrence plot-based approach for parkinson's disease identification. **Future Generation Computer Systems**, Elsevier, v. 94, p. 282–292, 2019. Disponível em: <<https://doi.org/10.1016/j.future.2018.11.054>>. Citado na página 42.
- AGARAP, A. F. Deep learning using rectified linear units (relu). **arXiv preprint arXiv:1803.08375**, 2018. Disponível em: <<https://doi.org/10.48550/arXiv.1803.08375>>. Citado na página 45.
- ALJAME, M.; AHMAD, I.; IMTIAZ, A.; MOHAMMED, A. Ensemble learning model for diagnosing covid-19 from routine blood tests. **Informatics in Medicine Unlocked**, Elsevier, v. 21, p. 100449, 2020. Disponível em: <<https://doi.org/10.1016/j.imu.2020.100449>>. Citado 2 vezes nas páginas 26 e 47.
- ALLAIN, C.; CLOITRE, M. Characterizing the lacunarity of random and deterministic fractal sets. **Physical review A**, APS, v. 44, n. 6, p. 3552, 1991. Disponível em: <<https://doi.org/10.1103/PhysRevA.44.3552>>. Citado na página 41.
- ALZUBAIDI, L.; ZHANG, J.; HUMAIDI, A. J.; AL-DUJAILI, A.; DUAN, Y.; AL-SHAMMA, O.; SANTAMARÍA, J.; FADHEL, M. A.; AL-AMIDIE, M.; FARHAN, L. Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions. **Journal of big Data**, Springer, v. 8, n. 1, p. 1–74, 2021. Disponível em: <<https://doi.org/10.1186/s40537-021-00444-8>>. Citado 3 vezes nas páginas 25, 26 e 28.
- ASSIS, T. A. d.; MIRANDA, J. G. V.; MOTA, F. d. B.; ANDRADE, R. F. S.; CASTILHO, C. M. C. d. Geometria fractal: propriedades e características de fractais ideais. **Revista Brasileira de Ensino de Física**, Sociedade Brasileira de Física, v. 30, n. 2, p. 2304.1–2304.10, 2008. ISSN 1806-1117. Disponível em: <<https://doi.org/10.1590/S1806-11172008000200005>>. Citado na página 19.
- BAIK, J.; YE, Q.; ZHANG, L.; POH, C.; ROSIN, M.; MACAULAY, C.; GUILLAUD, M. Automated classification of oral premalignant lesions using image cytometry and random forests-based algorithms. **Cellular Oncology**, Springer, v. 37, n. 3, p. 193–202, 2014. Disponível em: <<https://doi.org/10.1007/s13402-014-0172-x>>. Citado na página 37.
- BREIMAN, L. Random forests. **Machine learning**, Springer, v. 45, n. 1, p. 5–32, 2001. Disponível em: <<https://doi.org/10.1023/A:1010933404324>>. Citado na página 18.

BROADBENT, S. R.; HAMMERSLEY, J. M. Percolation processes: I. crystals and mazes. In: CAMBRIDGE UNIVERSITY PRESS. **Mathematical proceedings of the Cambridge philosophical society**. 1957. v. 53, n. 3, p. 629–641. Disponível em: <<https://doi.org/10.1017/S0305004100032680>>. Citado na página 24.

CANTRELL, C. D. **Modern mathematical methods for physicists and engineers**. Cambridge University Press, 2000. Disponível em: <<https://doi.org/10.1017/9780511811487>>. Citado na página 41.

CELANO, G. G. A resnet-50-based convolutional neural network model for language id identification from speech recordings. In: **Proceedings of the Third Workshop on Computational Typology and Multilingual NLP**. [s.n.], 2021. p. 136–144. Disponível em: <<http://dx.doi.org/10.18653/v1/2021.sigtyp-1.13>>. Citado na página 43.

DAS, A.; MOHANTY, M. N.; MALLICK, P. K.; TIWARI, P.; MUHAMMAD, K.; ZHU, H. Breast cancer detection using an ensemble deep learning method. **Biomedical Signal Processing and Control**, Elsevier, v. 70, p. 103009, 2021. Disponível em: <<https://doi.org/10.1016/j.bspc.2021.103009>>. Citado na página 47.

DENG, J.; DONG, W.; SOCHER, R.; LI, L.-J.; LI, K.; FEI-FEI, L. Imagenet: A large-scale hierarchical image database. In: **2009 IEEE Conference on Computer Vision and Pattern Recognition**. [s.n.], 2009. p. 248–255. Disponível em: <<https://doi.org/10.1109/CVPR.2009.5206848>>. Citado na página 43.

ECKMANN, J.-P.; KAMPHORST, S. O.; RUELLE, D. et al. Recurrence plots of dynamical systems. **World Scientific Series on Nonlinear Science Series A**, WORLD SCIENTIFIC PUBLISHING, v. 16, p. 441–446, 1995. Disponível em: <https://doi.org/10.1142/9789812833709_0030>. Citado na página 42.

ETZ, A. Introduction to the concept of likelihood and its applications. **Advances in Methods and Practices in Psychological Science**, Sage Publications Sage CA: Los Angeles, CA, v. 1, n. 1, p. 60–69, 2018. Disponível em: <<https://doi.org/10.1177/2515245917744314>>. Citado na página 31.

FARIA, P. R. de; CHAMMAS, R.; MELO, T. L. de; HSU, D. K.; LIU, F.-T.; NONOGAKI, S.; CARDOSO, S. V.; LOYOLA, A. M. Absence of galectin-3 does not affect the development of experimental tongue carcinomas in mice. **Experimental and molecular pathology**, Elsevier, v. 90, n. 2, p. 189–193, 2011. Disponível em: <<https://doi.org/10.1016/j.yexmp.2010.12.007>>. Citado na página 40.

FLORINDO, J. B. Descritores fractais aplicados à análise de texturas. 2013. Disponível em: <<https://doi.org/10.11606/T.76.2013.tde-02052013-161100>>. Citado na página 20.

GENTLE, J. E.; HÄRDLE, W. K.; MORI, Y. Springer handbooks of computational statistics. Springer, p. 985–1022, 2012. Disponível em: <<https://doi.org/10.1007/978-3-642-21551-3>>. Citado 2 vezes nas páginas 34 e 35.

GUPTA, S.; MOHANTA, S.; CHAKRABORTY, M.; GHOSH, S. Quantum machine learning-using quantum computation in artificial intelligence and deep neural networks: Quantum computation and machine learning in artificial intelligence. In: **2017 8th Annual Industrial Automation and Electromechanical**

Engineering Conference (IEMECON). [s.n.], 2017. p. 268–274. Disponível em: <<https://doi.org/10.1109/IEMECON.2017.8079602>>. Citado na página 53.

HALALLI, B.; MAKANDAR, A. Computer aided diagnosis - medical image analysis techniques. In: KUZMIAK, C. M. (Ed.). **Breast Imaging**. Rijeka: IntechOpen, 2017. cap. 5. Disponível em: <<https://doi.org/10.5772/intechopen.69792>>. Citado na página 17.

HIEU, N. V.; HIEN, N. L. H. Recognition of plant species using deep convolutional feature extraction. **Int J Emerg Technol**, 2020. Disponível em: <https://www.researchgate.net/publication/342334669_Recognition_of_Plant_Species_using_Deep_Convolutional_Feature_Extraction>. Citado na página 47.

HOSHEN, J.; BERRY, M.; MINSER, K. Percolation and cluster structure parameters: The enhanced hoshen-kopelman algorithm. **Physical Review E**, APS, v. 56, n. 2, p. 1455, 1997. Disponível em: <<https://doi.org/10.1103/PhysRevE.56.1455>>. Citado na página 25.

HOWARD, A. G.; ZHU, M.; CHEN, B.; KALENICHENKO, D.; WANG, W.; WEYAND, T.; ANDREETTO, M.; ADAM, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. **arXiv preprint arXiv:1704.04861**, 2017. Disponível em: <<https://doi.org/10.48550/arXiv.1704.04861>>. Citado 2 vezes nas páginas 43 e 44.

HUTTUNEN, M. J.; HRISTU, R.; DUMITRU, A.; FLOROIU, I.; COSTACHE, M.; STANCIU, S. G. Multiphoton microscopy of the dermoepidermal junction and automated identification of dysplastic tissues with deep learning. **Biomed. Opt. Express**, Optica Publishing Group, v. 11, n. 1, p. 186–199, Jan 2020. Disponível em: <<https://opg.optica.org/boe/abstract.cfm?URI=boe-11-1-186>>. Citado na página 37.

IVANOVICI, M.; RICHARD, N. Fractal dimension of color fractal images. **IEEE Transactions on Image Processing**, IEEE, v. 20, n. 1, p. 227–235, 2010. Disponível em: <<https://doi.org/10.1109/TIP.2010.2059032>>. Citado 2 vezes nas páginas 22 e 42.

IVANOVICI, M.; RICHARD, N.; DECEAN, H. Fractal dimension and lacunarity of psoriatic lesions-a colour approach. **medicine**, v. 6, n. 4, p. 7, 2009. Disponível em: <https://www.researchgate.net/publication/228806709_Fractal_Dimension_and_Lacunarity_of_Psoriatic_Lesions-A_Colour_Approach->. Citado 4 vezes nas páginas 20, 21, 23 e 24.

KADEMANI, D. **Improving Outcomes in Oral Cancer: A Clinical and Translational Update**. Springer Nature, 2019. Disponível em: <<https://doi.org/10.1007/978-3-030-30094-4>>. Citado 3 vezes nas páginas 5, 15 e 16.

KATHAMUTHU, N. D.; SUBRAMANIAM, S.; LE, Q. H.; MUTHUSAMY, S.; PANCHAL, H.; SUNDARARAJAN, S. C. M.; ALRUBAIE, A. J.; ZAHRA, M. M. A. A deep transfer learning-based convolution neural network model for covid-19 detection using computed tomography scan images for medical applications. **Advances in Engineering Software**, Elsevier, v. 175, p. 103317, 2023. Disponível em: <<https://doi.org/10.1007/978-3-030-30094-4>>. Citado 2 vezes nas páginas 42 e 43.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. **arXiv preprint arXiv:1412.6980**, 2014. Disponível em: <<https://doi.org/10.48550/arXiv.1412.6980>>. Citado na página 45.

- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F.; BURGESS, C.; BOTTOU, L.; WEINBERGER, K. (Ed.). **Advances in Neural Information Processing Systems**. Curran Associates, Inc., 2012. v. 25. Disponível em: <<https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>>. Citado 2 vezes nas páginas 5 e 28.
- KUMAR, V.; ABBAS, A. K.; FAUSTO, N.; MITCHELL, R. N. **Robbins patologia básica**. [S.l.]: Elsevier Brasil, 2008. Citado na página 15.
- LE, Q. V.; RANZATO, M.; MONGA, R.; DEVIN, M.; CHEN, K.; CORRADO, G. S.; DEAN, J.; NG, A. Y. **Building high-level features using large scale unsupervised learning**. 2012. Disponível em: <<https://doi.org/10.1109/ICASSP.2013.6639343>>. Citado na página 13.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, p. 436–44, 05 2015. Disponível em: <<https://doi.org/10.1038/nature14539>>. Citado na página 13.
- LOPES, R.; BETROUNI, N. Fractal and multifractal analysis: A review. **Medical image analysis**, Elsevier, v. 13, n. 4, p. 634–649, 2009. Disponível em: <<https://doi.org/10.1016/j.media.2009.05.003>>. Citado na página 22.
- MANDELBROT, B. B. Les objets fractals: forme, hasard et dimension. **(No Title)**, 1975. Citado na página 18.
- _____. The fractal geometry of nature. **San Francisco**, 2015. Citado 3 vezes nas páginas 5, 22 e 23.
- MANDELBROT, B. B.; FRAME, M. Fractals. **Hasard et Finances. 1st Edn., Flammarion, Paris**, 1997. Citado na página 19.
- MEDJAHED, S. A. A comparative study of feature extraction methods in images classification. **International journal of image, graphics and signal processing**, Modern Education and Computer Science Press, v. 7, n. 3, p. 16, 2015. Disponível em: <<https://doi.org/10.5815/ijigsp.2015.03.03>>. Citado na página 19.
- MEHANNA, H. M.; RATTAY, T.; SMITH, J.; MCCONKEY, C. C. Treatment and follow-up of oral dysplasia—a systematic review and meta-analysis. **Head & Neck: Journal for the Sciences and Specialties of the Head and Neck**, Wiley Online Library, v. 31, n. 12, p. 1600–1609, 2009. Disponível em: <<https://doi.org/10.1002/hed.21131>>. Citado na página 16.
- MOHANAIHAH, P.; SATHYANARAYANA, P.; GURUKUMAR, L. Image texture feature extraction using glcm approach. **International journal of scientific and research publications**, Citeseer, v. 3, n. 5, p. 1–5, 2013. Disponível em: <<https://www.ijsrp.org/research-paper-0513/ijsrp-p1750.pdf>>. Citado na página 13.
- MUMUNI, A.; MUMUNI, F. Data augmentation: A comprehensive survey of modern approaches. **Array**, v. 16, p. 100258, 2022. ISSN 2590-0056. Disponível em: <<https://doi.org/10.1016/j.array.2022.100258>>. Citado na página 32.
- NIKOLAIDIS, N. S.; NIKOLAIDIS, I. N.; TSOUROU, C. C. **A Variation of the Box-Counting Algorithm Applied to Colour Images**. 2011. Disponível em: <<https://doi.org/10.48550/arXiv.1107.2336>>. Citado na página 20.

PAUL, P.; PENNELL, M. L.; LEMESHOW, S. Standardizing the power of the hosmer–lemeshow goodness of fit test in large data sets. **Statistics in medicine**, Wiley Online Library, v. 32, n. 1, p. 67–80, 2013. Disponível em: <<https://doi.org/10.1002/sim.5525>>. Citado na página 31.

PEREZ, L.; WANG, J. The effectiveness of data augmentation in image classification using deep learning. **arXiv preprint arXiv:1712.04621**, 2017. Disponível em: <<https://doi.org/10.48550/arXiv.1712.04621>>. Citado na página 32.

PHIPHITPHATPHAISIT, S.; SURINTA, O. Food image classification with improved mobilenet architecture and data augmentation. In: . [s.n.], 2020. Disponível em: <<https://doi.org/10.1145/3388176.3388179>>. Citado 2 vezes nas páginas 6 e 44.

POHL, K. M.; FISHER, J.; BOUIX, S.; SHENTON, M.; MCCARLEY, R. W.; GRIMSON, W. E. L.; KIKINIS, R.; WELLS, W. M. Using the logarithm of odds to define a vector space on probabilistic atlases. **Medical Image Analysis**, Elsevier, v. 11, n. 5, p. 465–477, 2007. Disponível em: <<https://doi.org/10.1016/j.media.2007.06.003>>. Citado na página 31.

RAMOS, R. de S. A enfermagem oncológica no enfrentamento da pandemia de covid-19: reflexões e recomendações para a prática de cuidado em oncologia. **Revista Brasileira de Cancerologia**, v. 66, n. TemaAtual, 2020. Disponível em: <<https://doi.org/10.32635/2176-9745.RBC.2020v66nTemaAtual.1007>>. Citado na página 15.

RIBEIRO, M. G.; NEVES, L. A.; NASCIMENTO, M. Z. do; ROBERTO, G. F.; MARTINS, A. S.; TOSTA, T. A. A. Classification of colorectal cancer based on the association of multidimensional and multiresolution features. **Expert Systems with Applications**, Elsevier, v. 120, p. 262–278, 2019. Disponível em: <<https://doi.org/10.1016/j.eswa.2018.11.034>>. Citado na página 41.

RICHARDSON, L. F. Statistics of deadly quarrels. (**No Title**), 1960. Citado na página 22.

ROBERTO, G. F.; LUMINI, A.; NEVES, L. A.; NASCIMENTO, M. Z. do. Fractal neural network: A new ensemble of fractal geometry and convolutional neural networks for the classification of histology images. **Expert Systems with Applications**, Elsevier, v. 166, p. 114103, 2021. Disponível em: <<https://doi.org/10.1016/j.eswa.2020.114103>>. Citado na página 42.

ROBERTO, G. F.; NEVES, L. A.; NASCIMENTO, M. Z.; TOSTA, T. A.; LONGO, L. C.; MARTINS, A. S.; FARIA, P. R. Features based on the percolation theory for quantification of non-hodgkin lymphomas. **Computers in biology and medicine**, Elsevier, v. 91, p. 135–147, 2017. Disponível em: <<https://doi.org/10.1016/j.compbiomed.2017.10.012>>. Citado na página 24.

ROBERTO, G. F. et al. Associação entre atributos manuais e aprendizado profundo baseada em geometria fractal para classificação de imagens histológicas. Universidade Federal de Uberlândia, 2021. Disponível em: <<http://doi.org/10.14393/ufu.te.2021.656>>. Citado 6 vezes nas páginas 5, 6, 21, 22, 24 e 43.

ROKACH, L.; MAIMON, O. Decision trees. In: _____. [s.n.], 2005. v. 6, p. 165–192. Disponível em: <https://doi.org/10.1007/0-387-25465-X_9>. Citado na página 18.

- RUBIANO, G. El conjunto de mandelbrot. **Boletín de Matemáticas**, v. 3, n. 1, p. 25–36, 1996. Disponível em: <<https://repositorio.unal.edu.co/handle/unal/44715>>. Citado na página 19.
- SANT'ANA, J. M. D. A.; CHAMMAS, R.; LIU, F.-T.; NONOGAKI, S.; CARDOSO, S. V.; LOYOLA, A. M.; FARIA, P. R. D. Activation of the wnt/beta-catenin signaling pathway during oral carcinogenesis process is not influenced by the absence of galectin-3 in mice. **Anticancer research**, International Institute of Anticancer Research, v. 31, n. 9, p. 2805–2811, 2011. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/21868523/>>. Citado na página 40.
- SANTOS, D. F. D. d. et al. Automated segmentation of tumor regions from oral histological whole slide images using fully convolutional neural networks. Universidade Federal de Uberlândia, 2022. Disponível em: <<http://doi.org/10.14393/ufu.te.2022.252>>. Citado na página 27.
- SARICA, A.; CERASA, A.; QUATTRONE, A. Random forest algorithm for the classification of neuroimaging data in alzheimer's disease: A systematic review. **Frontiers in Aging Neuroscience**, v. 9, p. 329, 2017. ISSN 1663-4365. Disponível em: <<https://doi.org/10.3389/fnagi.2017.00329>>. Citado na página 30.
- SFERRAZZA, P. Grey level co-occurrence matrix and learning algorithms to quantify and classify use-wear on experimental flint tools. **Journal of Archaeological Science: Reports**, v. 48, p. 103869, 2023. ISSN 2352-409X. Disponível em: <<https://doi.org/10.1016/j.jasrep.2023.103869>>. Citado na página 18.
- SILVA, A.; ROZENDO, G.; TOSTA, T.; MARTINS, S.; LOYOLA, A.; CARDOSO, S.; LUMINI, A.; NEVES, L.; FARIA, P. de; NASCIMENTO, M. Zanchetta do. Cnn ensembles for nuclei segmentation on histological images of oed. In: . [s.n.], 2023. p. 601–604. Disponível em: <<https://doi.org/10.1109/CBMS58004.2023.00286>>. Citado 2 vezes nas páginas 26 e 40.
- SILVA, A. B.; MARTINS, A. S.; TOSTA, T. A. A.; NEVES, L. A.; SERVATO, J. P. S.; ARAÚJO, M. S. de; FARIA, P. R. de; NASCIMENTO, M. Z. do. Computational analysis of histological images from hematoxylin and eosin-stained oral epithelial dysplasia tissue sections. **Expert Systems with Applications**, Elsevier, v. 193, p. 116456, 2022. Disponível em: <<https://doi.org/10.1016/j.eswa.2021.116456>>. Citado 3 vezes nas páginas 37, 40 e 52.
- SILVA, A. B.; OLIVEIRA, C. I. D.; PEREIRA, D. C.; TOSTA, T. A.; MARTINS, A. S.; LOYOLA, A. M.; CARDOSO, S. V.; FARIA, P. R. D.; NEVES, L. A.; NASCIMENTO, M. Z. D. Assessment of the association of deep features with a polynomial algorithm for automated oral epithelial dysplasia grading. In: IEEE. **2022 35th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)**. 2022. v. 1, p. 264–269. Disponível em: <<https://doi.org/10.1109/SIBGRAPI55357.2022.9991758>>. Citado na página 52.
- SILVA, A. B. et al. Métodos computacionais para análise e classificação de displasias em imagens da cavidade bucal. Universidade Federal de Uberlândia, 2019. Disponível em: <<http://dx.doi.org/10.14393/ufu.di.2019.2390>>. Citado 3 vezes nas páginas 6, 40 e 41.

- SMITH, J.; RATTAY, T.; MCCONKEY, C.; HELLIWELL, T.; MEHANNA, H. Biomarkers in dysplasia of the oral cavity: a systematic review. **Oral oncology**, Elsevier, v. 45, n. 8, p. 647–653, 2009. Disponível em: <<https://doi.org/10.1016/j.oraloncology.2009.02.006>>. Citado 2 vezes nas páginas 12 e 16.
- SOILLE, P. Morphological texture analysis: An introduction. **Lecture Notes in Physics**, v. 600, 12 2002. Disponível em: <https://doi.org/10.1007/3-540-45782-8_9>. Citado 2 vezes nas páginas 18 e 19.
- SOUSA, W. B. d. Estimadores de máxima verossimilhança: casos que não satisfazem as condições de regularidade. 2018. Disponível em: <https://bdm.unb.br/bitstream/10483/24425/1/2018_WellingtonBernardoDeSousa_tcc.pdf>. Citado na página 31.
- TANG, X.-H.; KNUDSEN, B.; BEMIS, D.; TICKOO, S.; GUDAS, L. J. Oral cavity and esophageal carcinogenesis modeled in carcinogen-treated mice. **Clinical Cancer Research**, AACR, v. 10, n. 1, p. 301–313, 2004. Disponível em: <<https://doi.org/10.1158/1078-0432.CCR-0999-3>>. Citado na página 40.
- TANG, Z.; LI, M.; WANG, X. Mapping tea plantations from vhr images using obia and convolutional neural networks. **Remote Sensing**, v. 12, p. 2935, 09 2020. Disponível em: <<https://doi.org/10.3390/rs12182935>>. Citado 2 vezes nas páginas 6 e 44.
- TIAN, Y.; ZHANG, Y. A comprehensive survey on regularization strategies in machine learning. **Information Fusion**, Elsevier, v. 80, p. 146–166, 2022. Disponível em: <<https://doi.org/10.1016/j.inffus.2021.11.005>>. Citado na página 34.
- TILAKARATNE, W. M.; JAYASOORIYA, P. R.; JAYASURIYA, N. S.; SILVA, R. K. D. Oral epithelial dysplasia: Causes, quantification, prognosis, and management challenges. **Periodontology 2000**, Wiley Online Library, v. 80, n. 1, p. 126–147, 2019. Disponível em: <<https://doi.org/10.1111/prd.12259>>. Citado na página 15.
- TOPÎRCEANU, A.; GROSSECK, G. Decision tree learning used for the classification of student archetypes in online courses. **Procedia Computer Science**, v. 112, p. 51–60, 2017. ISSN 1877-0509. Disponível em: <<https://doi.org/10.1016/j.procs.2017.08.021>>. Citado na página 29.
- WARNAKULASURIYA, S.; REIBEL, J.; BOUQUOT, J.; DABELSTEEN, E. Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement. **Journal of oral pathology & medicine**, Wiley Online Library, v. 37, n. 3, p. 127–133, 2008. Disponível em: <<https://doi.org/10.1111/j.1600-0714.2007.00584.x>>. Citado 2 vezes nas páginas 12 e 16.
- WONG, D. D. E.; FUGLSANG, S. A.; HJORTKJÆR, J.; CEOLINI, E.; SLANEY, M.; CHEVEIGNÉ, A. de. A comparison of regularization methods in forward and backward models for auditory attention decoding. **Frontiers in Neuroscience**, v. 12, p. 531, 2018. ISSN 1662-453X. Disponível em: <<https://doi.org/10.3389/fnins.2018.00531>>. Citado na página 33.
- YAMASHITA, R.; NISHIO, M.; DO, R. K. G.; TOGASHI, K. Convolutional neural networks: an overview and application in radiology. **Insights into imaging**, Springer, v. 9, n. 4, p. 611–629, 2018. Disponível em: <<https://doi.org/10.1007/s13244-018-0639-9>>. Citado 4 vezes nas páginas 18, 26, 27 e 28.

YU, Y.; WANG, C.; FU, Q.; KOU, R.; HUANG, F.; YANG, B.; YANG, T.; GAO, M. Techniques and challenges of image segmentation: A review. **Electronics**, v. 12, n. 5, 2023. ISSN 2079-9292. Disponível em: <<https://doi.org/10.3390/electronics12051199>>. Citado na página 18.

ZHANG, Z.; SABUNCU, M. Generalized cross entropy loss for training deep neural networks with noisy labels. **Advances in neural information processing systems**, v. 31, 2018. Disponível em: <<https://doi.org/10.48550/arXiv.1805.07836>>. Citado na página 45.