
Classificação de Displasia da Cavidade Oral baseada em Descritores Fractais e *Ensemble learning*

Rafael Henrique de Oliveira Carvalho



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia
2024

Rafael Henrique de Oliveira Carvalho

**Classificação de Displasia da Cavidade Oral
baseada em Descritores Fractais e *Ensemble
learning***

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Ciência de Dados

Orientador: Marcelo Zanchetta do Nascimento

Uberlândia
2024

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

C331c
2024 Carvalho, Rafael Henrique de Oliveira, 1996-
Classificação de displasia da cavidade oral baseada em Descritores
Fractais e Ensemble learning [recurso eletrônico] / Rafael Henrique de
Oliveira Carvalho. - 2024.

Orientador: Marcelo Zanchetta do Nascimento.
Dissertação (Mestrado) - Universidade Federal de Uberlândia,
Programa de Pós-graduação em Ciência da Computação.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.di.2024.5077>
Inclui bibliografia.
Inclui ilustrações.

1. Computação. I. Nascimento, Marcelo Zanchetta do, 1976-,
(Orient.). II. Universidade Federal de Uberlândia. Programa de Pós-
graduação em Ciência da Computação. III. Título.

CDU: 681.3

André Carlos Francisco
Bibliotecário Documentalista - CRB-6/3408



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Ciência da
Computação

Av. João Naves de Ávila, 2121, Bloco 1A, Sala 243 - Bairro Santa Mônica, Uberlândia-MG,
CEP 38400-902

Telefone: (34) 3239-4470 - www.ppgco.facom.ufu.br - cpgfacom@ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação				
Defesa de:	Dissertação de Mestrado, 15/2024, PPGCO				
Data:	05 de maio de 2024	Hora de início:	14:00	Hora de encerramento:	15:20
Matrícula do Discente:	12122CCP015				
Nome do Discente:	Rafael Henrique de Oliveira Carvalho				
Título do Trabalho:	Classificação de Displasia da Cavidade Oral baseada em Descritores Fractais e Ensemble learning				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Ciência de Dados				
Projeto de Pesquisa de vinculação:	CPNq-311404/2021-9 FAPEMIG-APQ-00578-18				

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Mauricio Cunha Escarpinati- FACOM/UFU, Francisco de Assis Zampirolli - CMCC/UFABC e Marcelo Zanchetta do Nascimento - FACOM/UFU, orientador do candidato.

Os examinadores participaram desde as seguintes localidades: Francisco de Assis Zampirolli - Santo André/SP . Os outros membros da banca e a aluno participaram da cidade de Uberlândia.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Marcelo Zanchetta do Nascimento, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu à 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 Programa.

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:

Aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.
O competente diploma será expedido após cumprimento dos demais requisitos,

conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

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 06/06/2024, às 14:39, 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 **Francisco de Assis Zampirolli, Usuário Externo**, em 06/06/2024, às 17:00, 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 **Mauricio Cunha Escarpinati, Professor(a) do Magistério Superior**, em 07/06/2024, às 10:31, 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 **5401478** e o código CRC **6BCED8B6**.

*Dedico este trabalho a minha família por todo suporte, principalmente a minha esposa
Marina que sempre me apoiou e nunca me deixou desistir.*

Agradecimentos

Agradeço ao meu orientador, Marcelo Zanchetta do Nascimento, pela paciência na orientação, pelo convívio, pelo apoio, pela compreensão e pela amizade que tornaram possível a conclusão deste projeto.

O orientador deste projeto agradece o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (Projetos: 313643/2021-0 e 311404/2021-9) e da Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG (Projetos APQ-00578-18 e APQ-01129-21) e Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP (Projeto 2022/03020-1).

“Enquanto não se escolhe, tudo permanece possível.”
(Mr. Nobody)

Resumo

As lesões da cavidade oral podem ser classificadas em diferentes graus pelo especialista. No entanto, devido a carga de trabalho e nível de experiência essa tarefa pode estar sujeita a subjetividade. Uma forma de auxiliar o especialista nessa tarefa nos últimos anos, tem sido o emprego de sistemas computacionais como ferramenta suplementar para tomada de decisão do especialista. Neste trabalho, é apresentado uma abordagem para classificar lesões da cavidade oral com base em descritores obtidos da geometria fractal, modelos de redes neurais convolucionais e um *ensemble learning*. Na primeira etapa, o algoritmo *gliding-box* foi aplicado para obtenção dos descritores locais da geometria fractal em uma análise em multiresolução. Esse vetor de atributos foi remodelado para uma matriz 2D usando representações baseadas em gráfico de recorrência, *markov transition field* e *self-similarity matrix*. As matrizes 2D foram utilizadas como entrada para as redes convolucionais ResNet-50 e EfficientNet. Nesse estudo, foi investigado o impacto da otimização das camadas totalmente conectadas dessas estruturas. Finalmente, os modelos *ensemble learning* dos classificadores com a regra da soma foram aplicados aos conjuntos de dados. A eficiência da metodologia proposta foi analisada em um conjunto de dados composto por 74 regiões de interesse em displasia epitelial oral categorizadas entre as classes saudável, leve, moderada e grave. O método proposto utilizou a técnica de validação cruzada e obteve taxas de acurácia de até 98,6%. Essa abordagem torna-se uma ferramenta complementar para aplicação na prática clínica, auxiliando o especialista na tomada de decisão em relação à classificação das lesões.

Palavras-chave: Comitês de classificadores. Displasia. Imagens Histológicas. Processamento Digital de Imagens. Redes Neurais Convolucionais.

Abstract

Lesions in the oral cavity can be classified into different grades by specialists. However, due to workload and level of experience, this task may be subject to subjectivity. One way to assist the specialist in this task in recent years has been the use of computational systems as supplementary tools for decision-making. In this work, an approach is presented to classify lesions in the oral cavity based on descriptors obtained from fractal geometry, models of convolutional neural networks, and ensemble learning. In the first step, the gliding-box algorithm was applied to obtain local descriptors of fractal geometry in a multiresolution analysis. This attribute vector was reshaped into a 2D matrix using representations based on recurrence plot, Markov transition field, and self-similarity matrix. The 2D matrices were used as input for the ResNet-50 and EfficientNet convolutional networks. In this study, the impact of optimizing the fully connected layers of these structures was investigated. Finally, ensemble learning models of classifiers with the sum rule were applied to the datasets. The efficiency of the proposed methodology was analyzed on a dataset composed of 74 regions of interest in oral epithelial dysplasia categorized into healthy, mild, moderate, and severe classes. The proposed method used cross-validation technique and achieved accuracy rates of up to 98.6%. This approach becomes a complementary tool for application in clinical practice, aiding the specialist in decision-making regarding lesion classification.

Keywords: Ensemble Learning. Dysplasia. Histological Images. Digital Image Processing. Convolutional Neural Networks.

Lista de ilustrações

Figura 1 – Estrutura Anatômica da Boca. Adaptado de: Montero e Patel (2015)	30
Figura 2 – Exemplos de tecidos histológicos da região da cavidade oral: (a) tecido saudável, (b) DEO leve, (c) DEO moderada e (d) DEO severa. As setas na cor vermelha indicam núcleos celulares e as setas na cor azul indicam o epitélio. Fonte: Silva <i>et al.</i> (2019).	32
Figura 3 – Sistema para análise e detecção de células em histopatologia. Adaptado de: (HE <i>et al.</i> , 2022)	34
Figura 4 – Imagem de um relâmpago sob duas escalas de observação diferentes: (a) imagem completa e (b) imagem em uma escala quatro vezes menor. Fonte: (ROBERTO, 2021)	36
Figura 5 – Estrutura fractal construída de forma iterativa. Fonte: (ASSIS; MIRANDA; MOTA; ANDRADE; CASTILHO, 2008)	36
Figura 6 – Comparação da dimensão euclidiana e dimensão fractal. Fonte: (GONSALVES, 2015)	37
Figura 7 – Aplicação do algoritmo <i>gliding-box</i> em uma imagem com dimensões 8×8 pixels. Fonte: Roberto (2021)	37
Figura 8 – Curva logarítmica de DF em relação tamanho da caixa. Fonte: Roberto (2021)	39
Figura 9 – Representação em a) da imagem original e b) a imagem com os aglomerados definidos. A imagem apresenta 15×15 pixels em que há percolação devido à existência de um aglomerado que se estende entre duas extremidades da matriz, representado na cor verde-claro. Fonte: Roberto (2021)	40
Figura 10 – Exemplos de neurônios: (a) neurônio biológico (fonte: (SILVA, 2019)) e (b) neurônio artificial (adaptada de Haykin, Haykin, Haykin, Elektroingenieur e Haykin (2009)).	44
Figura 11 – Modelo de uma rede <i>Perceptron</i> Adaptado de: Rosebrock (2017)	44

Figura 12 – Exemplo da operação de convolução em uma imagem com um kernel de tamanho 3×3 <i>pixels</i> . Adaptado de: (YAMASHITA; NISHIO; DO; TOGASHI, 2018)	46
Figura 13 – Exemplos de um CNN com as diversas camadas. Adaptado de: Gurucharan (2020)	46
Figura 14 – Exemplos de funções de ativação (Step, Sigmoid, tanh, ReLU e suas variações). Adaptado de: Rosebrock (2017)	47
Figura 15 – Exemplos da operação de um <i>Max-Pooling</i> . Fonte: (SAVYAKHOSLA, 2019)	47
Figura 16 – Modelo simplificado das camadas da ResNet-50. Fonte: (BENDJILLALI; BELADGHAM; MERIT; TALEB-AHMED, 2020).	49
Figura 17 – Modelo simplificado das camadas da EfficientNet. Fonte: (AHMED; SABAB, 2022).	49
Figura 18 – Representação dos tipos de transferência de aprendizado: (a) baseado em instância, (b) baseado em mapeamento, (c) baseado em rede e (d) baseado em contradição. Adaptado de: Tan <i>et al.</i> (2018)	50
Figura 19 – Comportamento de um conjunto de classificadores composto por três classificadores. Adaptado de: (POLIKAR, 2012)	51
Figura 20 – Abordagens para utilizar comitê de classificadores, (a) <i>bagging</i> e (b) <i>boosting</i> . Adaptado de: Gaurav (2020)	52
Figura 21 – Fluxograma das etapas de um AG. Fonte: (ABBASGHOLIPOUR; OMID; KEYHANI; MOHTASEBI, 2011).	54
Figura 22 – Visão geral da metodologia proposta baseada em representação de atributos fractais para classificação de displasia de tecidos da cavidade oral. Fonte: Elaborado pelo autor.	67
Figura 23 – Exemplos de tecidos histológicos da cavidade bucal: (a) saudável, (b) displasia leve, (c) displasia moderada e (d) displasia severa. Fonte: (SILVA <i>et al.</i> , 2019)	69
Figura 24 – Visão detalhada da metodologia proposta baseada em representação de atributos fractais. Fonte: Elaborado pelo autor.	69
Figura 25 – Exemplo de metodologia proposta ilustrando a separando em cada canal de cor RGB: (a) Representação pelo GR (b) Representação pelo MTF e (c) Representação pela SSM. Fonte: Elaborado pelo autor. . .	70
Figura 26 – Avaliação do <i>fold</i> 1 com representação GR durante o treinamento e validação da rede ResNet-50 <i>baseline</i> . Fonte: Elaborado pelo autor. . .	76
Figura 27 – Desempenho do <i>fold</i> 1 com representação GR durante o treinamento e validação da rede EfficientNet <i>baseline</i> . Fonte: Elaborado pelo autor. .	77
Figura 28 – Matriz de confusão para EfficientNet com as representações investigadas: A) O, B) GR, C) MTF e D) SSM, Fonte: Elaborado pelo autor. .	78

Figura 29 – Matriz de confusão da rede ResNet-50 para cada uma das representações investigadas: A) O, B) GR, C) MTF e D) SSM. Fonte: Elaborado pelo autor.	79
Figura 30 – Médias marginais estimadas para os modelos <i>baseline</i> , em que gráfico de dispersão é a estimativa do teste estatístico. Fonte: Elaborado pelo autor.	80
Figura 31 – Valores de médias marginais estimadas para as estruturas das CNNs em que gráfico de dispersão é a estimativa do teste estatístico. Fonte: Elaborado pelo autor.	81
Figura 32 – Matrizes de confusão da rede otimizada EfficientNet para as representações investigadas: (A) O, (B) GR, (C) MTF e (D) SSM. Fonte: Elaborado pelo autor.	82
Figura 33 – Matrizes de confusão da rede otimizada ResNet-50 para as representações investigadas: (A) O, (B) GR, (C) MTF e (D) SSM. Fonte: Elaborado pelo autor.	83
Figura 34 – Histograma da distribuição do número de neurônios na primeira camada totalmente conectada, em cor laranja, para a EfficientNet, e em cor azul para a Resnet-50. Fonte: Elaborado pelo autor.	88
Figura 35 – Médias marginais estimadas para as estruturas dos comitês investigados, em que gráfico de dispersão é a estimativa do teste estatístico. Fonte: Elaborado pelo autor.	90

Lista de tabelas

Tabela 1	– Alguns critérios usados para diagnosticar a displasia. Fonte: (WARNAKULASURIYA; REIBEL; BOUQUOT; DABELSTEEN, 2008). . .	31
Tabela 2	– Resultados da métrica acurácia para as diversas abordagens de entrada para classificação com as redes convolucionais EfficientNet e ResNet-50.	76
Tabela 3	– Resultados dos ρ -valores do teste <i>post – hoc</i> Tuckey para as representações investigadas independente da estrutura da rede.	78
Tabela 4	– Valores médios obtidos com as métricas Acurácia (Acc), Sensibilidade (Sens) e Especificidade (Esp) para cada uma das classes de imagens investigadas da cavidade oral com a arquitetura EfficientNet.	80
Tabela 5	– Valores médios obtidos com as métricas acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para cada uma das classes de imagens investigadas da cavidade oral com a arquitetura ResNet-50.	81
Tabela 6	– Valores de acurácia para as arquiteturas com otimização das camadas totalmente conectadas.	82
Tabela 7	– Resultados experimentais de acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para as classes investigadas com otimização das camadas totalmente conectadas com estrutura baseada em ResNet-50. . .	83
Tabela 8	– Resultados experimentais com valores da acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para as classes investigadas com otimização das camadas totalmente conectadas com estrutura baseada em EfficientNet.	84
Tabela 9	– Resultados de ρ -valor do teste <i>post hoc</i> para as representações investigadas comparando os valores não otimizados e otimizados, independente da estrutura da rede.	84
Tabela 10	– Avaliação do modelo EfficientNet com a otimização das camadas totalmente conectadas com uso do AG.	86
Tabela 11	– Resultados da otimização das camadas totalmente conectadas com AG para o modelo ResNet-50.	87

Tabela 12 – Desempenho dos modelos propostos com comitês das CNNs e representações extraídas de descritores fractais sem o processo de otimização das camadas totalmente conectadas	89
Tabela 13 – Resultados dos ρ -valores do teste <i>post-hoc</i> para os comitês investigados independente da estrutura da rede.	91
Tabela 14 – Resultados experimentais de acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para as classes investigadas com o modelos de comitês e estrutura baseada em ResNet-50.	92
Tabela 15 – Resultados experimentais de acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para as classes investigadas com o modelos de comitês e estrutura baseada em EfficientNet.	93
Tabela 16 – Resultados experimentais de Acc com uso de comitê de classificadores e com otimização para as redes investigadas.	94
Tabela 17 – Resultados dos ρ -valores do teste <i>post – hoc</i> para os comitês investigados independente da estrutura da rede.	94
Tabela 18 – Análise da metodologia proposta em relação às abordagens existentes para classificação de imagens orais presentes na literatura.	95

Lista de siglas

AG Algoritmo Genético

CAD *Computer Aided Diagnosis*

DF Dimensão Fractal

FN falso negativo

FP falso positivo

GR Gráfico de Recorrência

H&E Hematoxilina e Eosina

LAC Lacunariedade

MTF *Markov Transition Field*

PERC Percolação

RNA Rede Neural Artificial

SSM *Self-Similarity Matrix*

VP verdadeiro positivo

VN verdadeiro negativo

Sumário

1	INTRODUÇÃO	23
1.1	Motivação	25
1.2	Objetivos e Desafios da Pesquisa	26
1.3	Hipóteses	26
1.4	Contribuições	27
1.5	Organização da Dissertação	27
2	FUNDAMENTAÇÃO TEÓRICA	29
2.1	Câncer	29
2.1.1	Cavidade Oral	29
2.1.2	Displasia	31
2.2	Histologia	32
2.3	Sistema de Apoio ao Diagnóstico	33
2.4	Geometria Fractal	35
2.4.1	Algoritmo <i>Gliding-Box</i>	37
2.4.2	Dimensão Fractal	38
2.4.3	Lacunaridade	39
2.4.4	Percolação	40
2.5	Modelos de Representação 2D das Características Fractais . . .	41
2.5.1	Gráfico de Recorrência	42
2.5.2	<i>Markov Transition Field</i>	42
2.5.3	<i>Self-Similarity Matrix</i>	43
2.6	Redes Neurais Artificiais	43
2.7	Redes Neurais Convolucionais	45
2.8	Transferência de Aprendizado	49
2.9	Comitê de Classificadores	51
2.10	Algoritmos Bio-inspirados	53
2.10.1	Algoritmo Genético	53

2.11	Considerações Finais	55
3	TRABALHOS RELACIONADOS: ESTADO DA ARTE	57
3.1	Detecção e Classificação de Displasias	57
3.2	Geometria Fractal	60
3.3	Algoritmos Bioinspirados para Otimização de Modelos de CNNs	62
3.4	Comitê de Classificadores	63
3.5	Considerações finais	65
4	METODOLOGIA	67
4.1	Introdução	67
4.2	Banco de Imagens	68
4.3	Atributos Fractais	68
4.4	<i>Reshapes</i> para Representação dos Descritores	70
4.5	Redes Neurais Convolucionais	71
4.6	Otimização de parâmetros das CNNs	71
4.7	Comitê de Classificadores	72
4.8	Crterios de Avaliação	73
5	EXPERIMENTOS E ANÁLISE DOS RESULTADOS	75
5.1	Avaliação dos Modelos de Redes Neurais Convolucionais	75
5.2	Arquiteturas com Otimização das Camadas Totalmente Conec- tadas	79
5.3	Comitê de Classificadores	87
5.3.1	Comitê com as Estruturas das CNNs	87
5.3.2	Comitê com Estrutura Otimizada das Camadas Totalmente Conectadas das CNNs	92
5.4	Uma Visão Geral do Método Proposto com Estado da Arte . .	94
6	CONCLUSÃO	97
6.1	Principais Contribuições	98
6.2	Trabalhos Futuros	98
6.3	Contribuição em Produção Bibliográfica	98
	REFERÊNCIAS	101

Introdução

A displasia oral é um precursor relativamente comum do câncer de boca ou carcinoma de células escamosas que em sua progressão para o câncer varia de 6% a 36% (SMITH; RATTAY; MCCONKEY; HELLIWELL; MEHANNA, 2009). Cerca de 95% dos cânceres da cavidade oral são carcinomas de células escamosas. Apesar dos inúmeros avanços no tratamento, esse tumor maligno epitelial agressivo está entre o mais comum tipo de neoplasia no mundo e apresenta uma taxa de sobrevida de longo prazo inferior a 50% (ABBAS; FAUSTO; MITCHELL; KUMAR, 2008).

O carcinoma de células escamosas está diretamente associado ao aparecimento de lesões potencialmente malignas em tecidos das regiões da boca (AXÉLL, 1987). Dentre as principais lesões, destaca-se as displasias, que são aquelas em que há alterações morfológicas. Elas se desenvolvem a partir da progressão de uma hiperplasia epitelial, passando por diferentes níveis que podem ser categorizadas em leve, moderada e severa. Essa lesão ainda pode progredir para um estágio mais avançado em que há uma invasão do tecido conjuntivo subjacente progredindo para o carcinoma *in situ* (SCULLY, 2014). A melhor estratégia para o tratamento e para que a lesão não se transforme em um tumor maligno é a predição da evolução. Embora existam fatores que possam ser preditivos para a transformação maligna dessas lesões bucais, a identificação da displasia e o grau da lesão são tarefas complexas.

Em histopatologia, um dos principais métodos para análise e investigação de tecidos com esse tipo de lesão se dá com o emprego de corantes como a hematoxilina e eosina, que colore os tecidos em tons de rosas e azuis. Estudos moleculares usando essas imagens têm auxiliado na predição da evolução do estado dessas lesões. Pelo fato das neoplasias não apresentarem padrão histológico e ter comportamento clínico diferenciado, a determinação de um diagnóstico preciso não é uma tarefa trivial (WRIGHT; ADDIS; LEONG, 2011). Normalmente, a discriminação entre os diferentes graus da lesão é complexa e repleta de subjetividade. Nessa tarefa, um especialista analisa com uso de um microscópio as características presentes na estrutura externa e interna dos núcleos e estruturas celulares presentes na amostra. Como o especialista possuiu uma grande carga de trabalho, essa

tarefa pode reduzir a acurácia de suas análises.

Nas últimas décadas, o uso dos microscópios de luz, equipados com câmeras digitais, auxiliam consideravelmente os patologistas na avaliação de lâminas dos tecidos histológicos. As imagens digitais obtidas por esses equipamentos permitem o desenvolvimento de algoritmos computacionais para processamento digital das imagens das amostras, podendo fornecer uma ferramenta para o sistema computacional de apoio ao diagnóstico (do inglês, *Computer Aided Diagnosis* (CAD)) (HE *et al.*, 2022).

Em sistemas CAD, é essencial o uso de descritores de características, cuja função pode ser sintetizada em transformar a imagem em um conjunto de dados que os represente. A etapa de extração de característica é muito importante para o sucesso da execução dos algoritmos de classificação. Existem diversos métodos disponíveis na literatura para extração de características como métodos baseados em modelos, dimensão fractal, entropia, matriz de coocorrência. Contudo, não é possível encontrar uma abordagem universal que forneça resultados relevantes para qualquer tipo de imagem. Sendo assim, faz-se necessário investigar o comportamento dos vários métodos nos diferentes tipos de aplicação. Nos últimos anos, o uso de características fractais, aquelas obtidas através de conceitos da geometria fractal, tem apresentado resultados relevantes na classificação de imagens histológicas como apresentado no trabalho de Roberto (2021).

A geometria fractal é um conceito apresentado para o estudo de formas que não podem ser definidas pela geometria euclidiana (MORRISON, 1975). A geometria fractal permite representar as propriedades de uma imagem por meio de descritores que apresentam invariância na sua forma à medida em que a sua escala é alterada, mantendo-se a sua estrutura idêntica à original. As principais propriedades que podem ser extraídas por meio da geometria fractal são a Dimensão Fractal (DF), a Lacunariedade (LAC) e a Percolação (PERC). A DF é uma medida que classifica a irregularidade e a complexidade de um fractal. A LAC indica como o espaço está preenchido. Enquanto que a PERC tem a capacidade de descrever as propriedades relacionadas à presença, à dimensão e à quantidade de aglomerados presentes nas imagens. Esses atributos vêm sendo investigados em imagens histopatológicas como nos trabalhos de Ivanovici, Richard e Decean (2009) e Roberto *et al.* (2017).

Os atributos fractais geram um vetor de características com alta dimensionalidade, uma vez que é possível obter atributos locais para uma região da imagem e globais para toda a imagem. Esses dados podem não permitir analisar redundância ou padrões que se repetem (ROBERTO; NEVES; LUMINI; MARTINS; NASCIMENTO, 2024). Dessa forma, novas estratégias vêm sendo desenvolvidas buscando transformar um vetor de atributos 1D em uma matriz 2D, para que possam trazer novas informações contribuindo para uma etapa de classificação. Existem diversos métodos de representação como: a representação sequencial de características (LUMINI; NANNI, 2018), o gráfico de recorrência (ECKMANN; KAMPHORST; RUELLE *et al.*, 1995), o *Self-Similarity Matrix*

(SSM) (FOOTE, 1999), o *Gramian Angular Field* (GAF) (WANG; OATES, 2015a) e o *Markov Transition Field* (MTF) (WANG; OATES, 2015b).

Nos últimos anos, os modelos de redes neurais convolucionais (do inglês, *convolutional neural network* - CNN) têm sido empregados com êxito nesse contexto. Com o uso de CNN, pode-se investigar a imagem de duas formas, seja a imagem do tecido ou a representação 2D dos atributos dessa imagem. A primeira forma explora a extração de características com a rede, que pode ser classificada, posteriormente, por outro tipo de classificador, como florestas aleatórias e *Support Vector Machine* (SVM). A segunda explora a classificação dos vetores de atributos nas próprias camadas totalmente conectadas. No entanto, o desempenho dos modelos estão sujeitos a diversos problemas, como o desbalanceamento de amostras de cada classe investigada assim como a maldição da dimensionalidade podem comprometer uma classificação (SAGI; ROKACH, 2018). Uma das técnicas que tem a objetivo de contornar esse problema é o comitê de classificadores.

Usar um comitê de classificadores pode ser uma solução com capacidade de treinar diferentes modelos e combinar suas predições, com resultados relevantes obtidos em histopatologia (KASSANI; KASSANI; WESOLOWSKI; SCHNEIDER; DETERS, 2019), (SAGI; ROKACH, 2018). Essa abordagem consiste em aplicar um procedimento para associar os resultados gerados por dois ou mais classificadores e, a partir dessa união, fornecer uma decisão tomada com base no comportamento dos classificadores empregados. Dessa forma, a combinação de abordagens de novas representações de atributos em conjunto com CNN e modelos de classificação mais robustos podem melhorar a detecção de lesões em imagens histológicas.

Portanto, é importante novos estudos para investigar e desenvolver metodologias que possam aprimorar a descrição de características dos objetos presentes nas imagens assim como associação com modelos de classificação mais robustos.

1.1 Motivação

Um sistema CAD é uma ferramenta poderosa que atua como leitor suplementar e pode ajudar inúmeros pacientes para que sejam diagnosticados de forma adequada. O diagnóstico precoce e a detecção do nível de displasias são muito importantes para que o paciente receba o tratamento adequado antes que a lesão evolua para um quadro mais grave, trazendo risco para a sua vida. Nesses sistemas, o uso de técnicas baseadas em geometria fractal têm apresentado resultados promissores na quantificação de imagens histológicas como nos trabalhos de Ivanovici, Richard e Decean (2009), Roberto *et al.* (2017) e Roberto (2021).

Apesar disso, ainda é possível explorar novas metodologias para uso desses atributos, por exemplo, ao associar técnicas de geometria fractal a uma arquitetura de rede CNN, visando melhorar os resultados obtidos em classificações. Várias representações de veto-

res em imagens 2D podem ser exploradas para melhorar ainda mais o desempenho do modelo. Usar um comitê de classificadores está entre as abordagens mais bem sucedidas em aprendizado de máquina e aplicações nos últimos anos, devido à sua capacidade de treinar diferentes modelos e combinar as predições (SAGI; ROKACH, 2018).

1.2 Objetivos e Desafios da Pesquisa

Este trabalho apresenta uma nova abordagem de classificação de displasia em imagens histológicas da cavidade oral baseado em representação 2D de características multiresolução fractal e associação de redes neurais convolucionais. Portanto, pontualmente, pretende-se:

- a) Investigar um método de extração de características baseados em descritores multiresolução de geometria fractal para imagens de displasias;
- b) Investigar as diferentes técnicas de representação de vetores 1D para matrizes 2D a fim de explorar novas representações dos dados;
- c) Desenvolver um comitê de modelos CNNs para melhorar a predição na detecção das lesões em imagens histopatológicas;
- d) Investigar um modelo otimizado de arquitetura CNN para o contexto de imagens de displasia bucal;
- e) Estudar o desempenho da metodologia proposta em comparação com trabalhos relacionados na literatura por meio de métricas de desempenho.

1.3 Hipóteses

O desenvolvimento de aplicações para análise de imagens histológicas tem apresentado resultados interessantes no estudo de lesões de displasias na cavidade bucal (ADEL *et al.*, 2018a). O uso de atributos fractais tem contribuído na descrição de informações presentes em tecidos histológicos (IVANOVICI; RICHARD; DECEAN, 2009), (ROBERTO *et al.*, 2017) e (ROBERTO, 2021). Entretanto, acredita-se que esses resultados podem ser ainda melhores, assim faz-se necessário buscar validar as seguintes afirmações nessa pesquisa: (a) O uso de atributos baseados em geometria fractal quando empregados em representação 2D e redes CNNs podem melhorar a classificação na determinação dos níveis de displasia em imagens da cavidade oral; (b) O uso de estratégias de otimização das camadas totalmente conectadas podem aumentar os valores de acurácia em modelos CNNs; (c) O emprego de comitês de modelos CNNs podem melhorar o desempenho da classificação na rotulação dos níveis de displasia; (d) A associação de modelos de representação e comitês de CNNs otimizados podem melhorar os valores de acurácia em relação aos modelos não otimizados.

1.4 Contribuições

As contribuições deste trabalho são:

- ❑ Um método computacional para classificação de imagens de tecidos histológicos da cavidade oral;
- ❑ Avaliação da eficiência do uso atributos fractais em modelo de método computacional para classificação de imagens de tecidos histológicos da cavidade oral;
- ❑ Investigação comparativa dos diferentes métodos de representação vetores de atributos 1D para matrizes 2D;
- ❑ Criação de um modelo de comitê de classificadores multiclasss para melhorar a predição das lesões;
- ❑ Desenvolvimento de um modelo otimizado de arquitetura CNN para o contexto de imagens de displasia bucal.

1.5 Organização da Dissertação

Este capítulo apresentou de forma sucinta os objetivos da metodologia proposta e a motivação para o desenvolvimento deste trabalho. Os próximos capítulos deste trabalho são organizados da seguinte maneira:

- ❑ **Capítulo 2 - Fundamentação Teórica:** apresenta conceitos de lesões pré-câncer e displasias e de processamento de imagens importantes para compreender este trabalho;
- ❑ **Capítulo 3 - Processamento de Imagens Histológicas Orais: Estado da Arte:** apresenta o estado da arte para processamento digital de imagens de câncer oral. São detalhados trabalhos focados em processamento de imagens de lesões cancerosas da região da cavidade oral;
- ❑ **Capítulo 4 - Metodologia:** apresenta o processo de criação do banco de imagens utilizado no trabalho, descreve as técnicas empregadas nas etapas de segmentação, pós processamento, extração de características e classificação, bem como as métricas utilizadas para a avaliação do método;
- ❑ **Capítulo 5 - Experimentos e Análise dos Resultados:** são apresentados os resultados obtidos em cada etapa da metodologia, resultados quantitativos e resultados obtidos por outros métodos presentes na literatura. São apresentadas discussões e comparações entre o método proposto e os demais;

- **Capítulo 6 - Conclusão:** apresenta as conclusões do trabalho, o potencial, as limitações e trabalhos futuros.

Fundamentação Teórica

Nesta seção, são apresentados os conceitos teóricos os quais foram divididos em câncer, displasia e histologia relacionados ao domínio de estudo. Em seguida, são apresentados as estratégias computacionais como o sistema de apoio ao diagnóstico, geometria fractal, representação 2D, redes neurais convolucionais, transferência de aprendizado, comitê de classificadores e algoritmos bio-inspirados.

2.1 Câncer

A Organização Mundial da Saúde (OMS) define o câncer como um conjunto de doenças caracterizadas pelo crescimento desordenado e descontrolado de células que invadem tecidos e órgãos, podendo se espalhar para outras partes do corpo, processo conhecido como metástase (OMS, 2023). As células cancerígenas, geralmente, formam massas de tecido, que interferem no funcionamento normal do organismo podendo levar a morte.

A mortalidade por câncer pode ser reduzida se os casos forem detectados e tratados precocemente (OPAS, 2020). A tarefa de identificar as lesões em estágios iniciais, antes da transformação em carcinoma, torna-se essencial para proporcionar um tratamento mais eficaz ao paciente. As lesões, frequentemente, associadas aos hábitos de vida prejudiciais, evidenciam a necessidade de vigilância constante e intervenção precoce, visando evitar a progressão para um estado maligno. Dentre diversas variações do câncer, o câncer de cavidade oral é uma condição séria que afeta milhões de pessoas em todo o mundo, o qual está fortemente associado ao tabagismo e ao consumo de álcool, sendo uma das principais causas de morte por câncer em todo o mundo (MONTERO; PATEL, 2015).

2.1.1 Cavidade Oral

A cavidade oral se estende desde a borda vermelha dos lábios até as papilas circunvaladas da língua inferiormente, a junção do palato duro e mole superiormente. Na Figura 1 é apresentado os vários sublocais anatômicos da cavidade oral: lábio, língua oral, as-

soalho da boca, mucosa bucal, gengiva superior e inferior, trígono retromolar e palato duro (MONTERO; PATEL, 2015).

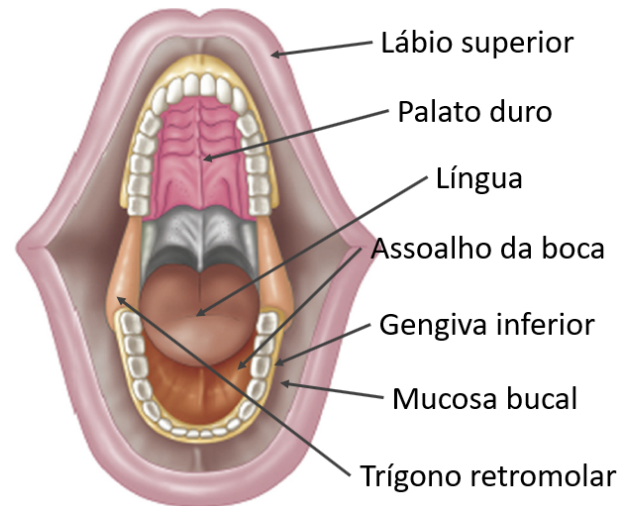


Figura 1 – Estrutura Anatômica da Boca. **Adaptado de:** Montero e Patel (2015)

As lesões da cavidade oral podem abranger uma ampla gama de condições, desde pequenas irritações até problemas mais graves, como a displasia oral. A displasia oral é caracterizada por alterações celulares anormais no revestimento da boca, que podem se desenvolver em lesões pré-cancerígenas se não forem tratadas. Contudo, apesar da criticidade da doença, um diagnóstico precoce pode favorecer um tratamento mais eficaz e humanizado da doença. Nesse contexto, a identificação de lesões pré-cancerosas torna-se essencial.

As lesões pré-cancerosas apresentam potencial maligno de tal forma que podem evoluir para um câncer (KUMAR; ABBAS; ASTER, 2013). É importante ressaltar que nem todas as lesões progridem para um câncer, mas por estimularem a proliferação regenerativa e exporem as células as condições, podem ocasionar inflamações, aumentando as chances da evolução para lesões malignas. As lesões pré-cancerosas, geralmente, apresentam similaridade entre células doentes e saudáveis e sua taxa de crescimento é reduzida, podendo levar anos para ocorrer uma transformação mais severa. No entanto, fatores externos podem induzi-las as transformações malignas (KUMAR; ABBAS; ASTER, 2013). Essas lesões são categorizadas como leucoplasia, eritroplasia, displasia epitelial e fibrose submucosa oral (TILAKARATNE; JAYASOORIYA; JAYASURIYA; SILVA, 2019). Os autores em Wolk e Lingen (2023) ressaltam que essas lesões são pré-malignas e há risco de evoluírem para um carcinoma oral de células escamosas. No entanto, na perspectiva diagnóstica, a definição de displasia é problemática por várias razões. Os critérios para diagnosticar e classificar a Displasia Epitelial Oral (DEO) não são uniformemente padronizados e estão abertos a uma vasta gama de interpretações.

2.1.2 Displasia

A DEO é um tipo de lesão pré-cancerosa importante, pois é uma anormalidade benigna que apresenta potencial para se tornar maligna. Essa lesão é um precursor relativamente comum do Carcinoma de Células Escamosas Oral (CCEO) e a progressão para o câncer tem um índice de ocorrência entre intervalos de 6% a 36% (SMITH; RATTAY; MCCONKEY; HELLIWELL; MEHANNA, 2009). Apesar dos avanços no tratamento, o tratamento e a detecção em estágio inicial são importantes, pois o CCEO é a sexta neoplasia mais comum no mundo e apresenta uma taxa de sobrevida inferiores a 50% (ABBAS; FAUSTO; MITCHELL; KUMAR, 2008).

As DEOs, quando presentes em tecidos histológicos, são caracterizadas por um conjunto de alterações celulares como o crescimento anormal e deformações na estrutura do núcleo, perda de uniformidade celular e hipercromasia podendo apresentar também desordenação no tecido epitelial (KUMAR; ASTER; ABBAS, 2010). No estudo proposto por Warnakulasuriya, Reibel, Bouquot e Dabelsteen (2008), os autores apresentam uma tabela (ver Tabela 1) que contém alguns critérios utilizados para classificar uma lesão como DEO. Os autores em Smith, Rattay, McConkey, Helliwell e Mehanna (2009) explicam que essas lesões são divididas em três classes ou graus de intensidades: leve, moderado ou grave, a Figura 2 apresenta exemplos dessas lesões. Uma variação pode ocorrer entre as classes de acordo com a anormalidade, sendo que quanto mais alta a intensidade da anormalidade mais grave é caracterizada a lesão.

Tabela 1 – Alguns critérios usados para diagnosticar a displasia. **Fonte:** (WARNAKULASURIYA; REIBEL; BOUQUOT; DABELSTEEN, 2008).

Arquitetura	Citologia
Estratificação epitelial irregular	Variação anormal no tamanho nuclear (anisonucleose)
Perda de polaridade das células basais	Variação anormal da forma nuclear
Cumes rete em forma de gota	Variação anormal no tamanho das células (anisocitose)
Maior número de figuras mitóticas	Variação anormal da forma celular (pleomorfismo celular)
Mitoses anormalmente superficiais	Aumento do tamanho nuclear

Em DEOs, as lesões com nível leve apresentam alterações celulares pouco intensas e comprometem apenas um terço da camada epitelial. As lesões DEOs do tipo severa são identificadas por alterações em mais de dois terços do epitélio, apresentando células com alterações intensas (WARNAKULASURIYA; REIBEL; BOUQUOT; DABELSTEEN, 2008). Por fim, a classe moderada de DEOs compromete entre um e dois terços do tecido epitelial e a intensidade das alterações precisa ser avaliada para classificar a lesão (SMITH; RATTAY; MCCONKEY; HELLIWELL; MEHANNA, 2009).

Em lesões pré-cancerosas, como a DEOs, geralmente, há grande similaridade entre células doentes e saudáveis, e sua taxa de crescimento é reduzida (SILVA *et al.*, 2019). No processo de diagnóstico, as lesões de mesmo tamanho podem ser classificadas em graus

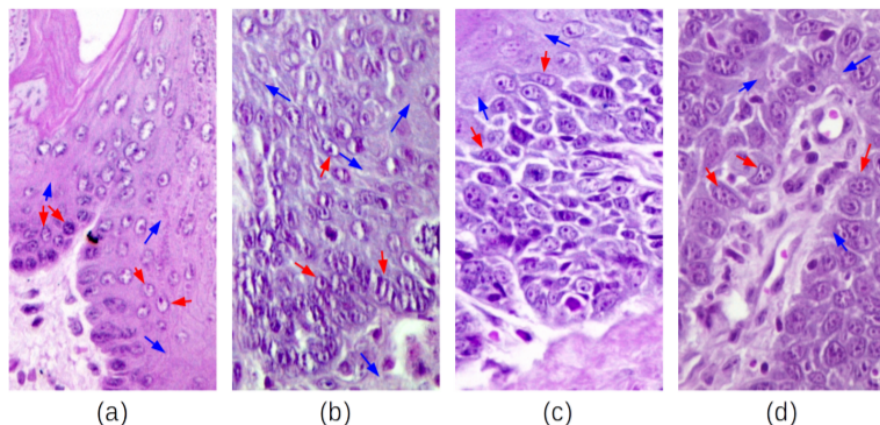


Figura 2 – Exemplos de tecidos histológicos da região da cavidade oral: (a) tecido saudável, (b) DEO leve, (c) DEO moderada e (d) DEO severa. As setas na cor vermelha indicam núcleos celulares e as setas na cor azul indicam o epitélio. **Fonte:** Silva *et al.* (2019).

diferentes pelo especialista, o que torna o diagnóstico subjetivo e dificulta a distinção entre as lesões (WARNAKULASURIYA; REIBEL; BOUQUOT; DABELSTEEN, 2008). A dificuldade de caracterização de padrões ainda pode levar a inconsistências no diagnóstico, provocando uma taxa de divergência variando entre 19% e 38% (MÜLLER, 2018). Com isso, é importante assegurar a confiabilidade e a qualidade das imagens utilizadas no diagnóstico. Durante o processo investigativo para o diagnóstico, uma amostra de tecido da cavidade oral é extraído e submetido a um processo de preparação, digitalização e captura com um microscópio óptico. Esse processo permite que o especialista possa explorar a diferenciação dos tecidos por uso de corantes, assim como a magnificação (aumento) da resolução, o que viabiliza a análise detalhada das estruturas.

2.2 Histologia

A histologia é o ramo da anatomia responsável pelo estudo dos tecidos e como são organizados para a formação dos órgãos constituintes do corpo. Trata-se da relação entre o modo no qual se encontra o tecido e as possíveis patologias associadas a ele. Por meio de imagens histológicas digitalizadas obtidas por um microscópio óptico, é possível estudar e encontrar padrões que auxiliam no diagnóstico e no tratamento de lesões na cavidade oral.

Para que a análise microscópica possa ser realizada, é necessário que o material seja fatiado em cortes histológicos, os quais são empregados na obtenção de lâminas, de forma a manter um aspecto similar ao seu estado natural (GARTNER, 2007). Para obter as lâminas, os autores em Junqueira e Carneiro (2013) explicam que é necessário uma etapa de fixação, na qual o tecido é imerso em soluções químicas com o objetivo de preservar a estrutura molecular e, em seguida, é feita a desidratação, em que o tecido é banhado em

álcool e xilol a fim de retirar a água do material, então, a operação de inclusão, em que o tecido é mergulhado em parafina para torná-lo mais rígido, conservando o material.

A microtomia, em que bloco de parafina do tecido é fatiado em secções uniformes de 1 a 10 micrômetros, é realizado após o processo de inclusão. Finalmente, ocorre a montagem e coloração, no qual o tecido é corado com Hematoxilina e Eosina (H&E) para permitir a identificação dos componentes. Após o processo da coloração, para proteger as estruturas, o tecido é colocado entre lâminas de vidro para análise em microscópio.

A hematoxilina, corante básico, possui um auxocromo catiônico (+) tendo assim, afinidade por componentes ácidos dos tecidos. O outro elemento, a eosina, corante ácido, possui um auxocromo aniônico (-), tendo assim afinidade por componentes básicos do tecido (catiônico) (CAPUTO; GITIRANA; MANSO, 2010). Assim, a hematoxilina cora os tecidos em tons de azul ou violeta as estruturas ácidas, como o núcleo e as porções dos citoplasmas ricas em ácido ribonucleico. Por outro lado, a eosina cora os componentes básicos celulares em tons róseos, como é o caso de vários componentes citoplasmáticos.

Após esse processo, os tecidos são digitalizados com o uso dos microscópios equipados com câmeras para aquisição de imagens digitais. Essas imagens são armazenadas em arquivos digitais e, geralmente, no modelo de cores RGB. Essas imagens digitalizadas são utilizadas para análise pelos especialistas, além de permitirem o desenvolvimento de técnicas de processamento de imagens para sistema de apoio aos profissionais numa avaliação das possíveis anormalidades presentes no tecido (GARTNER, 2007).

2.3 Sistema de Apoio ao Diagnóstico

O CAD, se tornou uma importante ferramenta empregada na rotina clínica no trabalho da detecção e classificação de diversas lesões presentes no corpo humano (HE *et al.*, 2022). Essas técnicas combinam o conhecimento de medicina e algoritmos computacionais para descobrir regras de diagnóstico baseadas em dados, simulando a tomada de decisões clínicas por um especialista, com objetivo de facilitar o diagnóstico para um tratamento adequado. Esses sistemas vêm sendo aplicados como uma solução suplementar para o especialista com objetivo de um diagnóstico mais efetivo com preços mais baratos (CHAN; HADJIISKI; SAMALA, 2020) e (YANASE; TRIANTAPHYLLOU, 2019).

A implementação de etapas para essa tarefa consiste em desenvolver algoritmos computacionais capazes de separar os objetos, extrair informações e assim classificá-las ao padrão das anormalidades. Um sistema CAD tem como base as fases: pré-processamento, segmentação, descrição e classificação (HE *et al.*, 2022). Na Figura 3, é apresentada as etapas dos processos aplicados as imagens histopatológicas.

A etapa de pré-processamento tem o intuito de melhorar a qualidade das características da imagem, acentuando-a e tornando-a mais adequada para análise ou aprimorando as informações para etapas posteriores. Na etapa de segmentação, a imagem passa por um

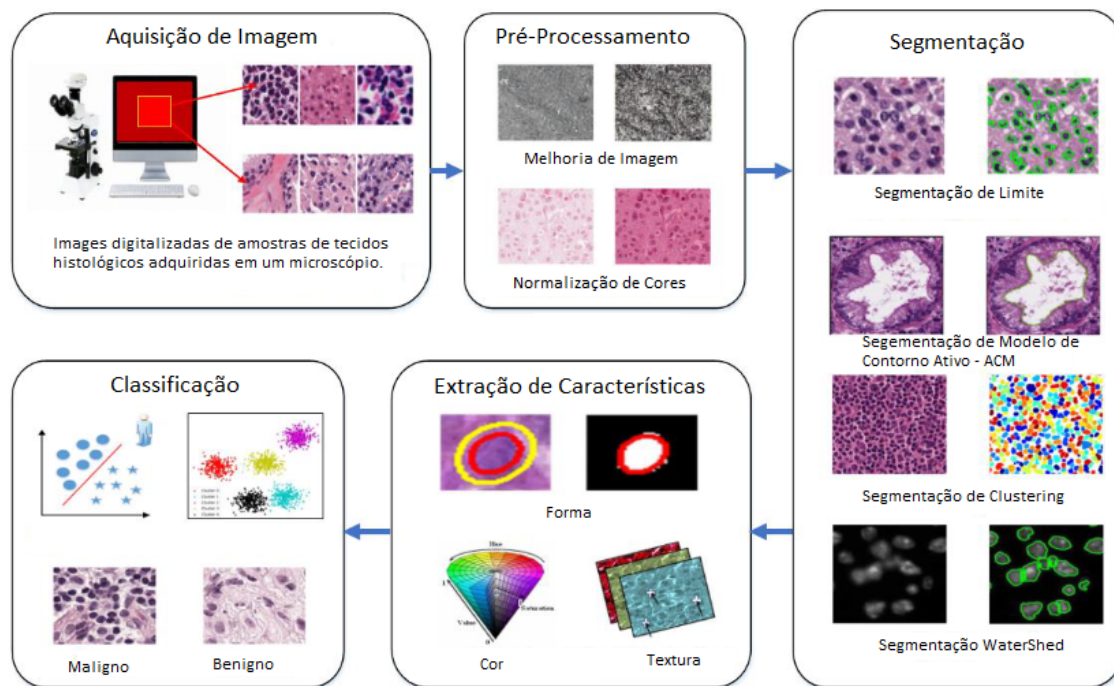


Figura 3 – Sistema para análise e detecção de células em histopatologia. **Adaptado de:** (HE *et al.*, 2022)

processo de identificação e separação das regiões de interesse ou objetos existentes. O reconhecimento e descrição refere-se a extração de características ou propriedades que discriminem as classes de objetos contidas na imagem.

De acordo com os autores em Jothi (2017), as imagens histológicas que representam o tecido maligno apresentam uma distribuição espacial mais irregular dos componentes celulares em comparação com o tecido saudável. Assim, as principais técnicas de extração de características aplicadas no contexto da histopatologia consistem em obter (ROBERTO, 2021):

- ❑ **Características morfológicas:** descrevem características simples, como a forma e o tamanho das estruturas celulares. Exemplos são os descritores de área, perímetro e concavidade;
- ❑ **Características topológicas:** descrevem a distribuição espacial dos componentes celulares. As principais incluem triangulação de *Delaunay*, diagrama de *Voronoi* e árvore geradora mínima;
- ❑ **Características de textura:** descrevem a variação de intensidade em diferentes regiões da imagem. Os principais métodos para obtenção de características de textura são dados pelas técnicas: matriz de co-ocorrência, filtro *Gabor*, transformada wavelet e geometria fractal (JOTHI, 2017).

A etapa de classificação é o processo de construir uma função ou projetar um modelo para mapear os dados extraídos de atributos para categorias específicas. Assim, é definida uma função ou modelo que mapeia as características para uma determinada categoria com objetivo de realizar a identificação da classe. Existem diversos algoritmos que apresentam resultados promissores, como árvores de decisão, *support vector machines* (SVMs), *K-nearest neighbors* (KNN) e rede artificial neural (HE *et al.*, 2022).

O uso de Inteligência Artificial (IA), principalmente, o aprendizado profundo (do inglês, *deep learning* - DL), em sistemas CAD, tem se popularizado e, cada vez mais, novos algoritmos são empregados para contribuir com soluções aos desafios presentes. Os autores em Roberto (2021) explicam que embora existam várias formas de quantificar uma imagem histológica, a geometria fractal é uma das abordagens que se destacou para a classificação de imagens histológicas, devido à semelhança existente entre o crescimento padrão de tumores e a arquitetura de um fractal (BAISH; JAIN, 2000) podendo ser uma abordagem promissora para representação e descrição das informações.

2.4 Geometria Fractal

A Geometria Fractal (GF) é um conceito apresentado para o estudo de formas que não pode ser definida pela geometria euclidiana, conforme descrito em Morrison (1975). Essa abordagem permite quantificar um objeto em relação a invariância na sua forma à medida em que a sua escala é alterada, mantendo-se a sua estrutura idêntica à original. Essa propriedade não ocorre, por exemplo, em uma circunferência, que parece reduzir a sua curvatura à medida em que amplia-se uma das suas partes. Na natureza, pode-se encontrar essas formas como no litoral de uma ilha, nas nuvens, árvores e raios.

Uma das principais propriedades que caracteriza a GF é a auto-similaridade, isto é, quando uma porção de uma figura ou de um contorno, pode ser vista como uma réplica do todo, em uma escala menor. Na Figura 4, é apresentada essa propriedade em que se observa a forma natural de um relâmpago, o qual não pode ser descrita pela geometria euclidiana clássica. Observando a imagem, na Figura, 4(a) nota-se a representação completa de um raio e uma representação (Figura 4(b)) tem a escala reduzida apresentando apenas uma parte da estrutura do raio. Essas imagens são similares com escalas distintas.

Ainda não existe uma definição formal para auto-similaridade, contudo, essa propriedade pode ser definida como características presentes na imagem, as quais podem ser observadas tanto em escalas menores, em subregiões, ou na imagem completa (BRACHMANN; REDIES, 2017).

Outra propriedade importante da GF é a complexidade infinita. O processo de geração de uma imagem fractal é recursivo. Portanto, quando é executado o processo para gerar a imagem, há um subprocedimento do próprio procedimento anteriormente executado. Essa etapa pode ser repetida infinitas vezes, permitindo obter uma estrutura infinitamente

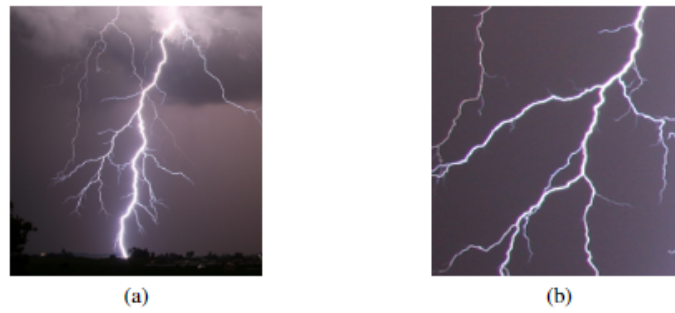


Figura 4 – Imagem de um relâmpago sob duas escalas de observação diferentes: (a) imagem completa e (b) imagem em uma escala quatro vezes menor. **Fonte:** (ROBERTO, 2021)

complexa (ASSIS; MIRANDA; MOTA; ANDRADE; CASTILHO, 2008). Na Figura 5, é apresentada essa propriedade para um estrutura fractal construída iterativamente.

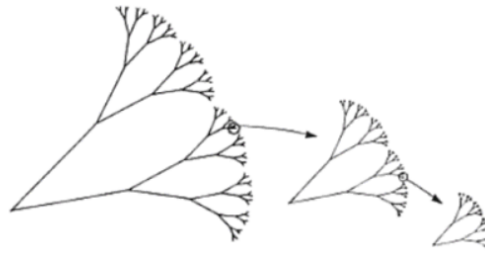


Figura 5 – Estrutura fractal construída de forma iterativa. **Fonte:** (ASSIS; MIRANDA; MOTA; ANDRADE; CASTILHO, 2008)

Além das propriedades discutidas, a Dimensão Fractal (DF) é definida como uma métrica que caracteriza padrões pela quantificação da complexidade relacionada a mudança dos detalhes para a mudança nas escalas. Essa métrica, ao contrário do que ocorre na geometria euclidiana, não é necessariamente um valor inteiro. Na Figura 6, é apresentada a diferença em que na dimensão euclidiana, uma estrutura representada por um ponto possui dimensão 0 (zero), uma linha possui dimensão 1 (um), uma superfície possui dimensão 2 (dois) e um volume possui dimensão 3 (três). No caso da DF, a representação é fracionária. Os autores em Roberto (2021) destacam que, em geral, um fractal deve ter detalhes em pequenas escalas, o que é incapaz de ser descrito pela geometria tradicional. Ainda, auto-similaridade aproximada ou estatística e definição possivelmente recursiva.

Em visão computacional, existem algumas abordagens para a observação das propriedades de auto-similaridade por meio da DF. Destaca-se os algoritmos *box counting* (NIKOLAIDIS; NIKOLAIDIS; TSOUROS, 2011) e *gliding-box* (IVANOVICI; RICHARD; DECEAN, 2009). Esses algoritmos propõem dividir as imagens em diferentes escalas e extrair características de cada subimagem, as quais resultam em descritores locais e globais como a DF, a Lacunariedade (LAC) e a Percolação (PERC) (LOPES; BETROUNI, 2009).

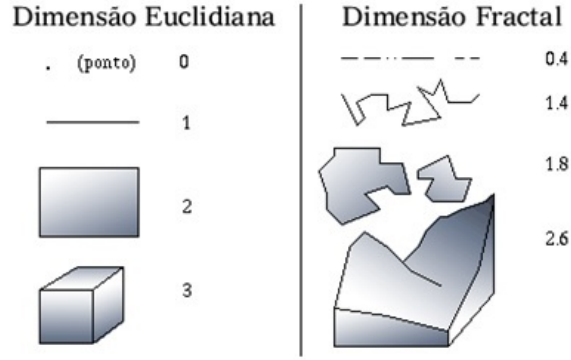


Figura 6 – Comparação da dimensão euclidiana e dimensão fractal. **Fonte:** (GONSALVES, 2015)

Para as imagens médicas coloridas, o modelo apresentado por Morrison (1975) tem limitações com as propriedades dos canais de cores, ou seja, não consegue tratar essas propriedades. Esse problema foi resolvido nos estudos de Ivanovici, Richard e Decean (2009) e (ROBERTO *et al.*, 2017). Assim, as características fractais podem ser obtidas de imagens coloridas explorando as propriedades desses algoritmos.

2.4.1 Algoritmo *Gliding-Box*

Os autores em Roberto (2021) descrevem que o algoritmo *gliding-box* consiste em posicionar uma caixa de tamanho $L \times L$ na parte superior do lado esquerdo da imagem para percorrer as informações, sendo L a dimensão da caixa em *pixels*. Essa caixa se move da esquerda para a direita da região superior para a região inferior da imagem. Após deslizar pela imagem, a caixa é reposicionada no ponto inicial e o valor de L é incrementado em 2 unidades, como ilustrado na Figura 7. O valor de L deve ser sempre um número ímpar, pois é necessário um pixel central para as próximas etapa do método. Em uma imagem de tamanho $H \times W$, o número T de caixas β_i em função de L é dado pela Equação 1:

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

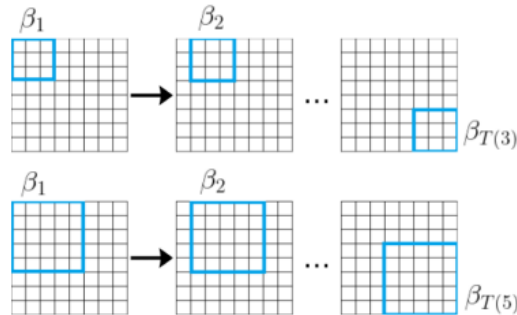


Figura 7 – Aplicação do algoritmo *gliding-box* em uma imagem com dimensões 8×8 pixels. **Fonte:** Roberto (2021)

Em cada caixa, é realizado uma análise de similaridade das cores para cada pixel. Essa análise é feita fixando-se o pixel central e atribuindo a um vetor $F_{ci} = (x_{ci}, y_{ci}, r_{ci}, g_{ci}, b_{ci})$, dado por x_{ci} e y_{ci} que correspondem às coordenadas espaciais do pixel na imagem. Os parâmetros r_c , g_c e b_c correspondem à intensidade luminosa de cada um dos canais de cores de uma imagem no formato RGB.

Os demais pixels da caixa, $F_i = (x_i, y_i, r_i, g_i, b_i)$, são comparados ao pixel central por meio de uma distância, sendo, por exemplo, a distância de Minkowski, conforme exibido na Equação 2. Cada pixel que satisfaça d , ou seja, sua distância até o pixel central é menor ou igual ao tamanho da caixa L , é considerado como pertencente a caixa. Portanto, se o valor de d correspondente à distância entre f_i e f_{ci} for menor ou igual a L então f_i , esse pixel é rotulado como 1, caso contrário, é rotulado como 0 (ROBERTO, 2021).

$$d = \max(|f_i - f_{ci}|) \leq L, \forall i = \overline{1, 5}. \quad (2)$$

Para a distância de Manhattan, deve ser empregado a Equação 3 em que p apresenta o valor 1 enquanto que para distância Euclidiana p apresenta o valor 2.

$$D(x, y) = (\sum |x_i - y_i|^p)_{i=1}^{\frac{1}{p}}. \quad (3)$$

Após calcular o número de pixels, que satisfaçam a Equação 2, para cada caixa e para diferentes valores de L , as informações obtidas são usadas para a construção de uma matriz de probabilidade. Nessa matriz, cada elemento corresponde à probabilidade de que m pixels em uma escala L sejam rotulados como 1. Por fim, a matriz é normalizada de modo que a soma de todos os elementos de uma coluna seja igual a 1. Essa matriz permite obter os valores locais e globais de algumas métricas fractais.

2.4.2 Dimensão Fractal

A métrica DF pode ser definido como uma métrica que caracteriza padrões pela quantificação de sua complexidade relacionadas as mudanças de detalhes para escalas (MANDELROT; MANDELROT, 1982). Assim, dado uma matriz de probabilidade, para cada tamanho de caixa L , e para uma escala definida como DF local, pode-se calcular a DF aplicando a Equação 5:

$$DF_{local} = \sum_{m=1}^N \frac{P(m, L)}{m}. \quad (4)$$

Após o cálculo para todos os valores de L , é possível construir um gráfico que relaciona a evolução de $\log L \times \log N(L)$. O valor do coeficiente angular da reta gerado corresponde ao valor global de DF. Na Figura 8, é ilustrado o cálculo da DF para uma imagem do banco de imagens do trabalho de Roberto (2021). É possível notar que nem sempre uma reta perfeita é gerada, principalmente, devido à presença de ruído, o que afeta a análise em escalas menores. Nesses casos, um método de regressão linear deve ser aplicado para a obtenção da característica.

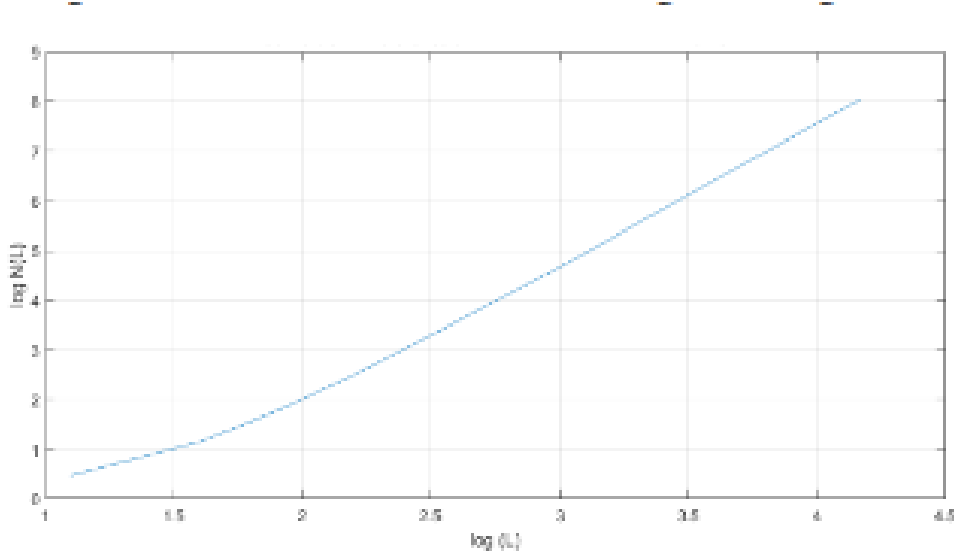


Figura 8 – Curva logarítmica de DF em relação tamanho da caixa. **Fonte:** Roberto (2021)

A DF pode ser definida como uma medida que avalia o quanto do espaço da imagem está preenchido (ROBERTO, 2021). Entretanto, uma limitação dessa medida ocorre quando imagens diferentes apresentam valores similares ou até mesmo idênticos da DF. Visando solucionar esse problema, outras medidas baseadas na GF são empregadas, como por exemplo, a lacunaridade e a percolação.

2.4.3 Lacunaridade

A LAC é uma medida complementar à DF que avalia de que forma o espaço de um fractal é preenchido. Assim como na medida DF, em seu trabalho (IVANOVICI; RICHARD; DECEAN, 2009), os autores propuseram um método para a obtenção dessa propriedade em imagens coloridas. Essa métrica é calculada a partir de uma matriz de probabilidade, sendo computado os momentos, de primeira e segunda ordem, respectivamente, a média e o desvio-padrão, usando as Equações 6 e 7. Os valores m e $P(m, L)$ são os mesmos utilizados na Equação 5.

$$\mu(L) = \sum_{m=1}^N mP(m, L). \quad (5)$$

$$\mu^2(L) = \sum_{m=1}^N m^2P(m, L). \quad (6)$$

Então, a medida LAC, em uma escala L , é dada por $\lambda(L)$, que pode ser obtida pela Equação 8.

$$\lambda(L) = \frac{\mu^2(L) - (\mu(L))^2}{(\mu(L))^2}. \quad (7)$$

Assim como a DF, com essa métrica, é possível calcular os valores globais. Esses atributos são obtidos através da curva de $\log(L) \times \log\lambda(L)$, calculando a área sob a

curva, obliquidade, razão de área e escala do ponto máximo (CĂLIMAN; IVANOVICI, 2012).

2.4.4 Percolação

A PERC é um fenômeno físico definido por (BROADBENT; HAMMERSLEY, 1957) referente ao transporte de fluidos em meios porosos. Diversos sistemas naturais apresentam esse tipo de comportamento, sendo o escoamento de água através de pó de café ou gases por um meio rochoso são bons exemplos desse fenômeno. Se um desses fluidos, dado pela água ou gás, é capaz de atravessar todo o sistema de uma extremidade a outra, diz-se que há percolação. Esses conceitos podem também ser adaptados para a área de análise de imagens (ROBERTO, 2021).

Para tanto, seja uma imagem como apresentado na Figura 9 em que alguns *pixels* da grade foram rotulados com pontos. Suponha-se que esses pixels correspondem aos poros, que na natureza são os espaços pela qual um fluido é capaz de atravessar sobre um meio. Então, por meio da Figura 9(b) se observa facilmente que os pontos formam vários aglomerados, cada um deles representados por uma cor. Em uma imagem, é dito que ocorre uma percolação se existe um aglomerado em que se estende de uma extremidade da imagem à qualquer outra. Assim, no exemplo da Figura 9, nota-se a ocorrência de uma percolação, considerando conectividade.

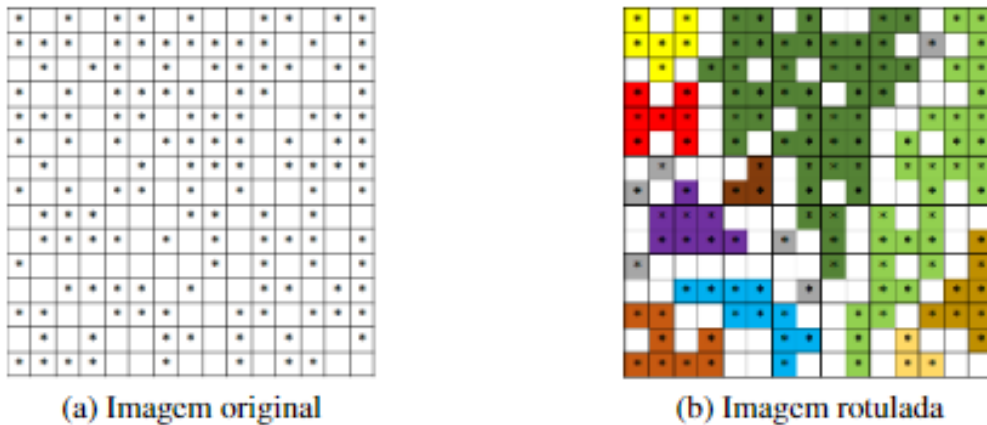


Figura 9 – Representação em a) da imagem original e b) a imagem com os aglomerados definidos. A imagem apresenta 15×15 pixels em que há percolação devido à existência de um aglomerado que se estende entre duas extremidades da matriz, representado na cor verde-claro. **Fonte:** Roberto (2021)

Em sistemas naturais, a distribuição dos poros ocorre de forma aleatória, assim existe em cada sistema um valor denominado p , que corresponde à probabilidade de um espaço ser ou não um poro (MALARZ, 2015). O valor p ainda possui uma importante propriedade conhecida como limiar de percolação. Essa propriedade denota que, a partir de um determinado valor de p , há a garantia da ocorrência de percolação. Esses conceitos são

adaptados para análise de imagens e evitam esforços nas buscas de aglomerados que indiquem percolação (ROBERTO, 2021). Cada tipo de sistema terá um valor diferente de limiar de percolação durante o processo de análise. Baseados nos trabalhos de Dean (1963), Ziff (1986) e Bird e Perrier (2010), para matrizes quadradas, que é o caso de imagens coloridas, o valor de limiar é de 0,59275.

Tendo esse valor de limiar, os autores em Ribeiro *et al.* (2018) extraíram atributos de percolação de uma imagem usando conceitos da geometria fractal. Nesse método, os *pixels* que satisfizeram a condição para a distância de Minkowski, foram definidos como poros e identificados por meio do algoritmo de rotulagem de Hoshen-Kopelman (HOSHEN; BERRY; MINSER, 1997). Assim, em cada escala L , foi extraído o número médio de aglomerados por caixa, a taxa de ocorrência de percolação e área média do maior aglomerado. A Equação 8 apresenta o cálculo do número médio de aglomerados, o qual é a soma do total de aglomerados C em cada caixa i dividido pelo número total de caixas T .

$$C(L) = \frac{\sum_{i=1}^T C_i}{T} \quad (8)$$

Na Equação 9, é mostrado como a taxa de ocorrência de percolação Q é calculada, assim, conforme citado anteriormente ele representa o número de pixels rotulados como poros que satisfazem o limite de percolação de 0,59275 na caixa i dividido pelo número total de caixas T .

$$Q = \frac{\sum_{i=1}^T q_i}{T} \quad (9)$$

Por fim, a Equação 10 corresponde ao cálculo da área de cobertura média do maior aglomerado M , dada pela soma das taxas de ocupação dos maiores aglomerados em cada caixa i dividido pelo número total de caixas T .

$$M(L) = \frac{\sum_{i=1}^T \frac{M_i}{L^2}}{T} \quad (10)$$

2.5 Modelos de Representação 2D das Características Fractais

Cada característica obtida pelas abordagens da GF pode ser calculada de forma global, para toda a imagem por meio da análise de um gráfico em função do tamanho da caixa, ou de forma local, isto é, para cada caixa. Como a quantidade de caixas depende do tamanho da imagem (Equação 1), os atributos obtidos pela GF podem gerar um vetor de características com alta dimensionalidade. Estudos têm buscado explorar novas abordagens para representação dos dados de forma a identificar as estruturas comprometidas numa avaliação de grupos de dados (AFONSO *et al.*, 2019). Assim, a construção de uma matriz de malhas, a partir de dados discretos, pode ser usada como uma represen-

tação para avaliação de descritores e classificação. Dentre as estratégias disponíveis na literatura, nessa seção, são apresentadas as abordagens empregadas neste trabalho.

2.5.1 Gráfico de Recorrência

Uma forma de representar um vetor de características como uma malha 2D é por meio do uso do Gráfico de Recorrência (GR). Em aplicações da área médica, esse método vem sendo usado para a detecção da doença de Parkinson (AFONSO *et al.*, 2019). Essa técnica foi proposta por (ECKMANN; KAMPHORST; RUELLE *et al.*, 1995) e fornece uma maneira de visualizar a natureza periódica do comportamento de mudança de estado através de um espaço de fase, projetando eventos repetidos em espaços 2D ou 3D. Assim, dada uma trajetória x_i de um determinado sistema, o gráfico de recorrência é definido pela Equação 12, no qual N é o número de estados, x_i presente na trajetória, m é a dimensão de imersão, ρ é o raio de vizinhança no ponto x_i , θ é a função de *Heaviside* e $\| \cdot \|$ é a norma de vizinhança (REIS; DOMINGUES; MACAU, 2018).

$$\vec{R}_{ij}^{mp} = \theta(\rho - \| \vec{x}_i - \vec{x}_j \|); \vec{x}_i \in R^m; i, j = 1, \dots, N. \quad (11)$$

Em imagem, a aplicação dessa técnica em um conjunto de dados pode ser empregada usando os N atributos, resultando como saída uma matriz quadrada $N \times N$, que reduz a Equação 11 para a Equação 12, em que x_i e y_i representam o i -ésimo e o j -ésimo do vetor de características e $\| x_i - x_j \|$ é a norma da distância euclidiana.

$$R_{ij} = \| x_i - x_j \| \forall i, j = 1, \dots, N. \quad (12)$$

2.5.2 Markov Transition Field

O MTF é uma ferramenta poderosa para representar a estrutura de dependência entre variáveis aleatórias. Essa métrica tem sido usada em vários domínios de aplicação, incluindo aprendizado de máquina e visão computacional conforme já destacado no estudo de Salakhutdinov (2009). Em Wang e Oates (2015b), os autores definem que o MTF é calculado através da seguinte matriz:

$$\begin{bmatrix} \omega_{ij}|x_1 \in q_i, x_1 \in q_j} & \cdots & \omega_{ij}|x_1 \in q_i, x_n \in q_j} \\ \omega_{ij}|x_2 \in q_i, x_2 \in q_j} & \cdots & \omega_{ij}|x_2 \in q_i, x_n \in q_j} \\ \vdots & & \vdots \\ \omega_{ij}|x_n \in q_i, x_1 \in q_j} & & \omega_{ij}|x_n \in q_i, x_n \in q_j} \end{bmatrix} \quad (13)$$

Nesse método, o conjunto de dados é dividido em *bins*, no qual os *bins* contendo os dados i e j , são q_i e q_j , e ω_{ij} na matriz é a probabilidade de transição de $q_i \rightarrow q_j$ (WANG; OATES,

2015b). O tamanho da matriz é definido pela quantidade de *bins*. Esse método permite a representação de dados em uma matriz 2D, a qual pode ser usada como entrada de redes neurais convolucionais (ESTEBESARI; RAJABI, 2020).

2.5.3 Self-Similarity Matrix

Uma outra abordagem para representação de características é o SSM introduzido por Foote (1999). Nesse método, dado dois vetores v_i e v_j , então, é calculado uma métrica de similaridade vetorial S sobre uma janela w , como descrito na Equação 14, em que a função s é descrita na Equação 15.

$$S_w(i, j) = \frac{1}{w} \sum_{k=0}^{w-1} s(i+k, j+k). \quad (14)$$

$$s(i, j) = \frac{v_i * v_j}{|v_i| |v_j|}. \quad (15)$$

O resultado da operação é uma matriz construída de forma em que cada pixel na localização i, j recebe valor de intensidade de brilho proporcional à sua medida de similaridade. Essa técnica foi empregada na representação dos dados em uma matriz para entrada de uma rede neural convolucional como descrito no trabalho de Cohen-Hadria e Peeters (2017).

2.6 Redes Neurais Artificiais

A Rede Neural Artificial (RNA) é uma classe de algoritmos de aprendizado de máquina especializados no reconhecimento de padrões, inspirado na estrutura, na forma como os neurônios se relacionam na função do cérebro humano e como ocorre o processamento da informação. Uma RNA é composta por várias camadas, das quais cada uma tem uma função específica no processo de análise e classificação das informações.

As RNAs são formadas a partir de um modelo artificial de neurônio e tentam simular as conexões entre eles para gerar conhecimento sobre um conjunto de dados (HAYKIN; HAYKIN; HAYKIN; ELEKTROINGENIEUR; HAYKIN, 2009). Na biologia, os neurônios possuem quatro componentes principais:

- ❑ Os dendritos, que recebem informações de outros neurônios;
- ❑ O corpo celular que processa as informações recebidas e gera uma nova informação;
- ❑ Os axônios, que propagam a nova informação para outros neurônios;
- ❑ As sinapses que representam as conexões entre os axônios e os dendritos de neurônios diferentes.

Os neurônios artificiais simulam esses elementos biológicos por meio de sinapses, as quais são responsáveis pela transmissão de dados entre eles. Há uma operação matemática para processar múltiplos os dados recebidos e uma função de ativação, que normaliza a informação produzida pelo neurônio em um valor finito, definindo um peso para esse estímulo (HAYKIN; HAYKIN; HAYKIN; ELEKTROINGENIEUR; HAYKIN, 2009). Na Figura 10 é apresentado um comparativo ilustrativo entre um neurônio real e um neurônio artificial.

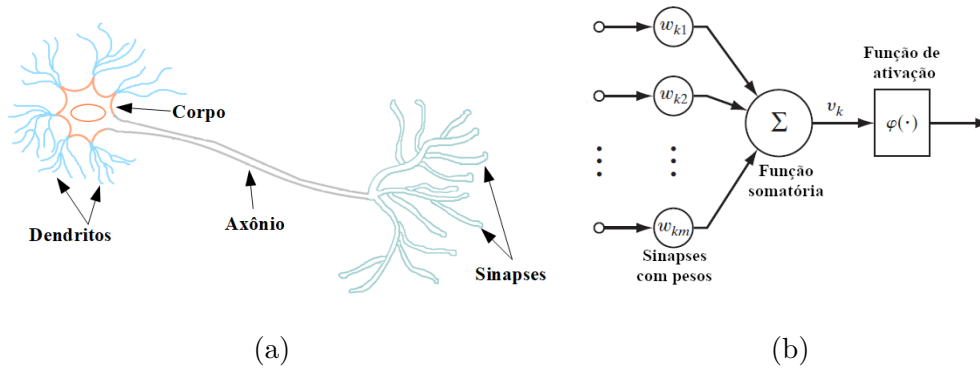


Figura 10 – Exemplos de neurônios: (a) neurônio biológico (fonte: (SILVA, 2019)) e (b) neurônio artificial (adaptada de Haykin, Haykin, Haykin, Elektroingenieur e Haykin (2009)).

A primeira estrutura de uma RNA surgiu em 1950, com o algoritmo denominado *Perceptron*, publicado por Rosenblatt Rosenblatt (1958). Esse modelo utilizou diversas entradas e somatórios ponderados para produzir como saída uma classificação. Na Figura 11, é apresentado um exemplo de uma rede *perceptron*. Nesse modelo, há um processo automático, em que os pesos dos neurônios são obtidos durante o processo de treinamento do algoritmo. Para um problema de classificação binária, por exemplo, as entradas são ponderadas e somadas passando por uma função de degrau unitário ou função de ativação, a qual permite definir a categoria (classe) dos dados.

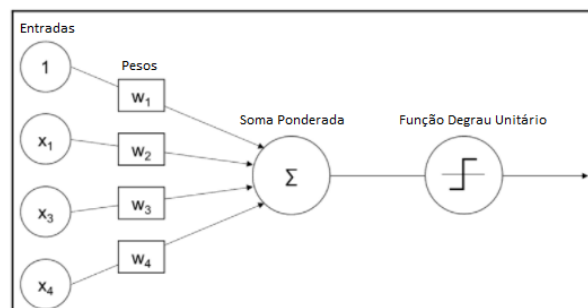


Figura 11 – Modelo de uma rede *Perceptron* Adaptado de: Rosebrock (2017)

Em Rosebrock (2017), o autor explica que esse método evoluiu ao longo dos anos, com os trabalhos de Rumelhart, Hinton e Williams (1986) e LeCun, Bottou, Orr e Müller (2012), e as investigações com algoritmos de retropropagação, pois se tinha uma RNA

feedforward de multicamadas. Isso permitiu melhorar as soluções considerando o aprendizado com os erros do processo. Atualmente, com o melhoramento das arquiteturas computacionais baseadas em placas de processamento gráfico (do inglês, *graphic processing unit* - (GPU)), essas abordagens são processadas em alta velocidade, permitindo novas pesquisas e contribuições para novas arquiteturas. Essas novas redes podem ser compostas com muito mais camadas ocultas, o que permite explorar aprendizagem hierárquica, em que conceitos simples são aprendidos nas camadas inferiores e mais padrões abstratos nas camadas superiores. Com isso, novos modelos de redes têm surgido para os diversos domínios de aplicação.

2.7 Redes Neurais Convolucionais

As Redes Neurais Convolucionais são arquiteturas empregadas como extratores e classificadores de imagem (LECUN; BOTTOU; BENGIO; HAFFNER, 1998). Vários trabalhos abordam o seu uso e maneiras de aprimorar as redes como os trabalhos de Hussain, Bird e Faria (2018) e Lei, Pan e Huang (2019). Rosebrock (2017) em seu trabalho, descreve que a grande diferença de uma CNN para as demais abordagens é que, ao invés de definir manualmente um conjunto de regras num classificador e ter métodos de extração de características, essas etapas são tratadas no modelo CNN durante o treinamento e os atributos são aprendidos para uma classificação com novos dados. Esse modelo permite unir um sistema capaz de extrair características e ao mesmo tempo fazer a classificação de dados baseados em um método supervisionado.

A CNN é um tipo especializado de rede neural para processamento de dados, que tem uma topologia semelhante a uma grade, a qual permite a exploração de série temporal (1-D) e de imagem (2-D), em que operações matemáticas dadas pela operação de convolução são empregadas nas etapas iniciais para extração de informações globais da imagem.

A convolução é uma operação linear que, a partir de duas funções dadas, resulta numa terceira que mede a soma do produto dessas funções ao longo da região subentendida pela superposição delas em função do deslocamento existente entre elas. Em um modelo de CNN, uma função pode corresponder a imagem, enquanto que a outra é o *kernel* aplicado no processo convolucional, o qual produzirá o mapa de características (GOODFELLOW; BENGIO; COURVILLE, 2016). Esse processo é demonstrado na Figura 12. No trabalho de (BIANCHI, 2020), o autor explica que durante a operação da convolução em uma imagem, o *kernel* vai se deslocando ao longo da imagem, como uma janela deslizante, em que operações de multiplicação e soma são aplicadas seguindo um deslocamento determinado pela operação de *stride*.

Na Figura 13, é ilustrado uma CNN padrão e como as diferentes camadas da rede estão conectadas. Existem diversas arquiteturas de CNNs com diferentes camadas que podem ser usadas em uma tarefa de classificação de imagens. As CNNs tradicionais

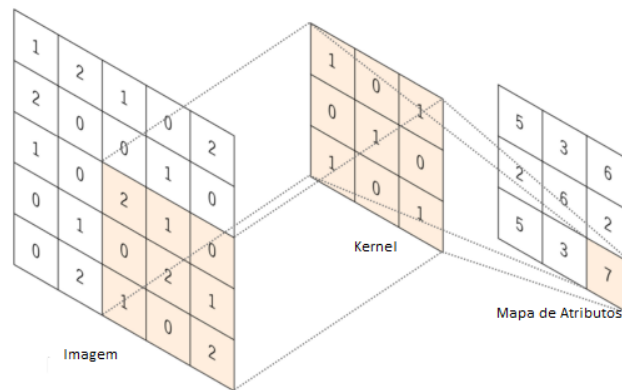


Figura 12 – Exemplo da operação de convolução em uma imagem com um kernel de tamanho 3×3 pixels. **Adaptado de:** (YAMASHITA; NISHIO; DO; TOGASHI, 2018)

apresentam, além da camada de convolução, a camada de ativação, o *pooling* e as camadas totalmente conectadas (do inglês, *fully-connected* - FC). A arquitetura da CNN é formada pela interação entre uma série dessas camadas de uma maneira específica (ROSEBROCK, 2017).

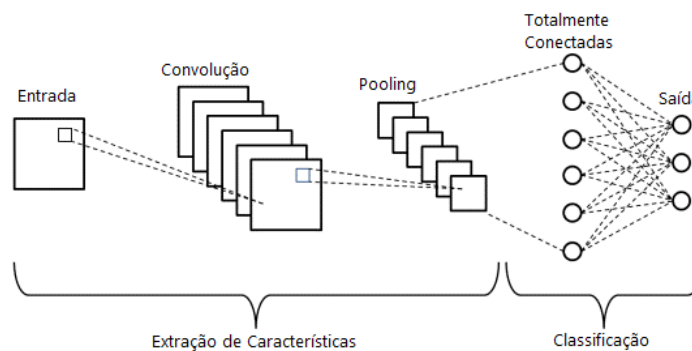


Figura 13 – Exemplos de um CNN com as diversas camadas. **Adaptado de:** Gurucharan (2020)

A camada de ativação é geralmente subsequente a uma camada convolucional, em que normalmente são aplicadas funções de ativação. Existem diversas funções de ativação, cada uma com suas vantagens e desvantagens. Dentre elas, destacam-se as funções de ativação binária, linear, sigmóide, tanh, *Rectified Linear Unit* - (ReLU) e Leaky ReLU ilustradas na Figura 14. A ReLU é a função mais frequentemente utilizada na camada de *pooling* em que sua principal característica é a não ativação de todos os neurônios ao mesmo tempo, assim, se a entrada for negativa, ela será convertida em zero e o neurônio não será ativado, o que torna a rede esparsa e eficiente para a computação (ROSEBROCK, 2017).

O *pooling* é uma camada responsável por reduzir progressivamente o tamanho do espaço de uma entrada. Isso permite reduzir a quantidade de parâmetros e uso computa-

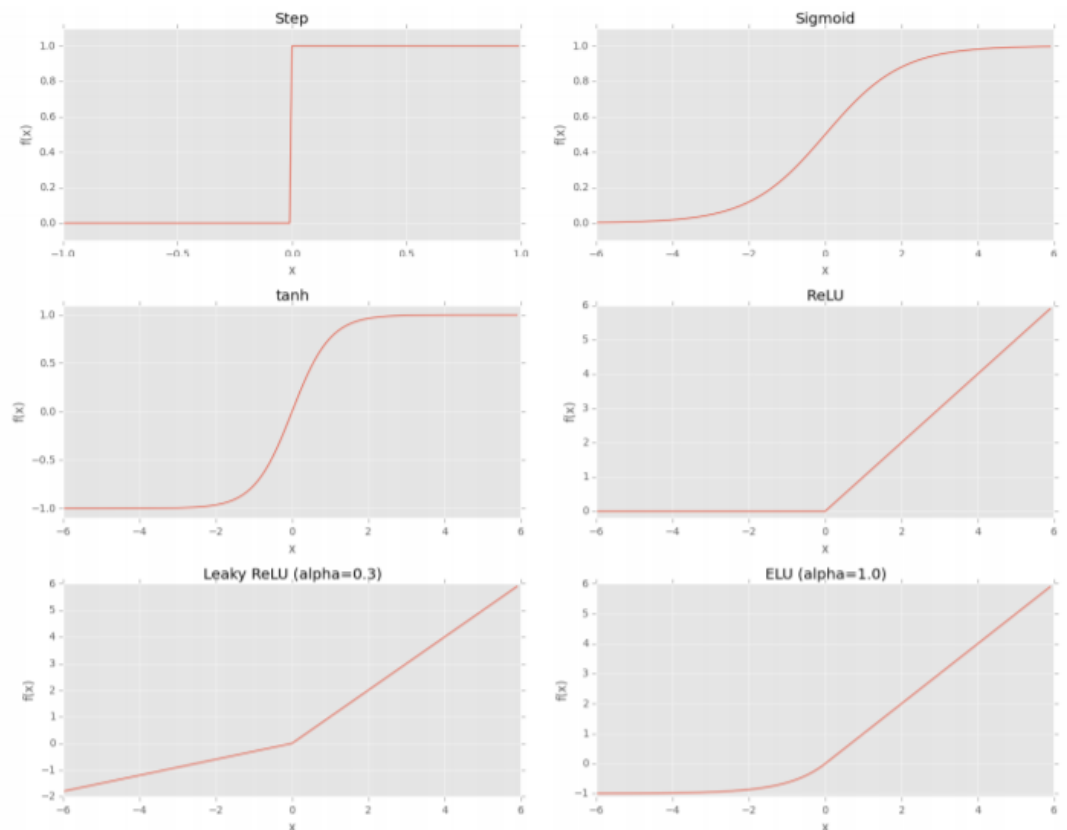


Figura 14 – Exemplos de funções de ativação (Step, Sigmoid, tanh, ReLU e suas variações). **Adaptado de:** Rosebrock (2017)

cional da rede, assim como minimizar os efeitos do *overfitting*. Geralmente, essa camada é inserida entre conjuntos consecutivos formados por camada convolucionais e de ativação. Uma ilustração desse processo pode ser vista na Figura 15, em que mostra uma técnica conhecida como *max-pooling*, no qual somente o máximo valor da janela é usado como saída da camada.

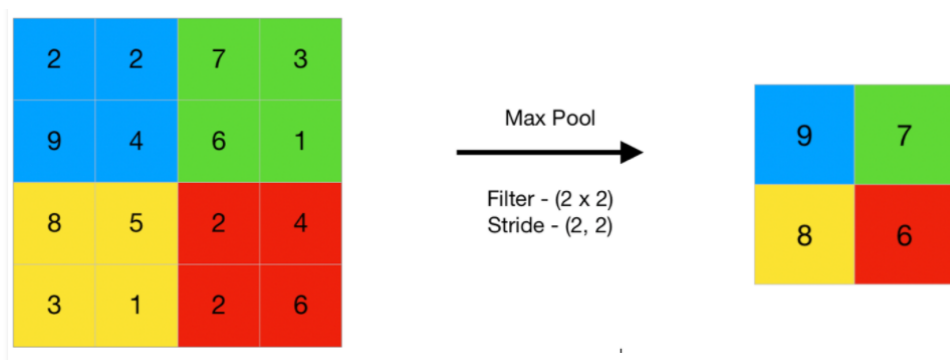


Figura 15 – Exemplos da operação de um *Max-Pooling*. **Fonte:** (SAVYAKHOSLA, 2019)

As camadas FCs fazem parte da etapa de classificação (ROSEBROCK, 2017). Essas camadas são conectadas a cada unidade de ativação da próxima camada. As FCs são conectadas por meio de um peso adaptado durante a fase de treinamento. O ajuste dos pesos geralmente é feito pelo uso do algoritmo retropropagação. Geralmente, o número

de *perceptrons* da última camada FC é igual ao número de classes da base de dados de treinamento (SINGH, 2018). Na camada totalmente conectada da rede, é comum encontrar também camada de *dropout*, cuja função é reduzir o sobreajuste do modelo. Na camada *dropout*, durante o treinamento, uma porção dos neurônios é aleatoriamente escolhida para ser temporariamente "desligada", o que significa que seus valores não são atualizados e suas contribuições são temporariamente retiradas da rede. Essa abordagem obriga a rede a aprender representações mais robustas e distribuídas, já que não pode depender exclusivamente de neurônios específicos durante o treinamento (SRIVASTAVA; HINTON; KRIZHEVSKY; SUTSKEVER; SALAKHUTDINOV, 2014).

Na literatura, diversos modelos CNN empregam essas diversas camadas com diversas formas de interação entre elas. Dentre eles, pode-se destacar a ResNet-50 (HE; ZHANG; REN; SUN, 2015) e EfficientNet (TAN; LE, 2019), modelos de aprendizado profundo que têm apresentado resultados relevantes na área de classificação de leões histológicas. A ResNet-50 é uma CNN com conceitos de redes residuais, ao adicionar blocos residuais ao modelo, permite que a rede aprenda "resíduos", isto é, a diferenças entre a entrada e a saída de um bloco, facilitando o treinamento de redes mais profundas, pois as ativações residuais podem ser propagadas por meio da rede de forma mais eficiente, evitando o problema do desaparecimento do gradiente. A ResNet-50 é composta por 50 camadas, incluindo camadas convolucionais, camadas de *pooling* e camadas totalmente conectadas. Por sua vez, a rede EfficientNet é uma CNN cujo principal vantagem é a maximização da eficiência computacional. A arquitetura EfficientNet é composta por um bloco de construção básico chamado *Mobile Inverted Bottleneck Convolution* (MBConv), que consiste em uma sequência de camadas de convolução, ativação e normalização, projetadas para maximizar a eficiência de cálculo.

Na Figura 16, é apresentado um modelo simplificado da estrutura Resnet-50. Esse modelo apresenta uma entrada padrão de entrada com uma imagem de 224x224 pixels e suas camadas convolucionais são compostas por filtros 3x3 e para as camadas que têm o mesmo número de filtros, o tamanho da saída só alterado quando o número de filtros é duplicado, dessa forma, o tamanho de saída é reduzido pela metade. Por fim, o modelo apresenta camada de *pooling* e uma camada totalmente conectada com *softmax*.

Na Figura 17, é apresentado um modelo simplificado da estrutura da EfficientNet. Esse modelo apresenta uma entrada padrão com uma imagem de 224x224 pixels. A rede começa com uma sequência de camadas convolucionais para extrair características básicas da imagem e segue uma sequência de Blocos Invertido de Convolução de *Bottleneck* (MBConv). Este bloco é composto por uma sequência de operações, começando com uma expansão linear da dimensionalidade, seguida por uma camada de convolução 3x3 e, finalmente, uma redução linear da dimensionalidade, que ajuda a reduzir a complexidade computacional. No final da rede, uma camada de média global é usada para produzir o vetor de características e a última camada é uma camada totalmente conectada com

softmax que mapeia as características para as classes de saída.

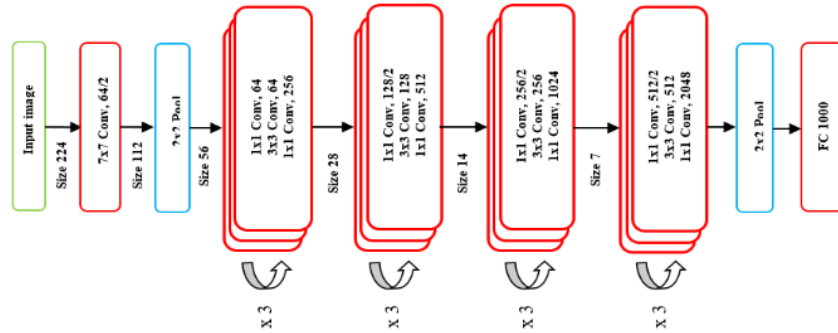


Figura 16 – Modelo simplificado das camadas da ResNet-50. **Fonte:** (BENDJILLALI; BELADGHAM; MERIT; TALEB-AHMED, 2020).

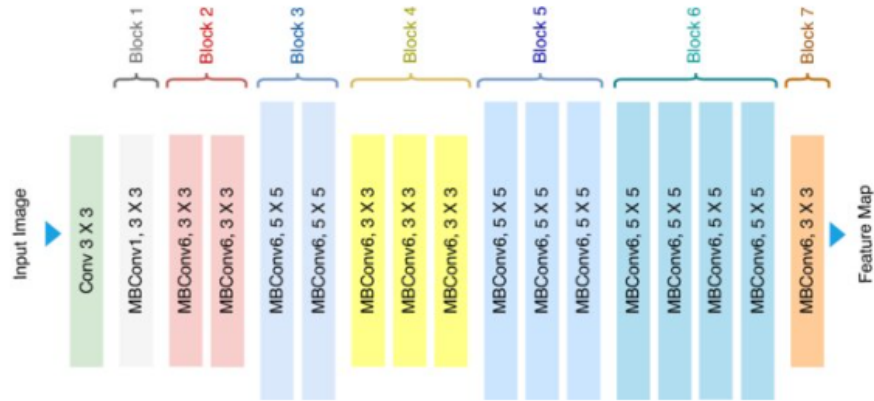


Figura 17 – Modelo simplificado das camadas da EfficientNet. **Fonte:** (AHMED; SABAB, 2022).

2.8 Transferência de Aprendizado

A transferência de aprendizado é um método muito conveniente e eficaz para treinar redes neurais visando reduzir o tempo gasto no treinamento, assim como tentar suprir a falta de grandes conjuntos de dados (TAN *et al.*, 2018). Essa técnica consiste em reciclar o conhecimento de modelos previamente treinados com grande volume de dados. As CNNs requerem GPUs e CPUs de alto desempenho para serem treinadas, porém com a transferência de aprendizado, elas podem ser implementadas em computadores pessoais, uma vez que o treinamento pode ser feito em tempo reduzido (LU; LU; ZHANG, 2019). Nessa técnica, ocorre primeiro treino em uma rede básica, em um conjunto de dados e tarefas básicas. Em seguida, há a transferência desses recursos aprendidos para uma outra rede para ser treinada em um conjunto de dados e tarefa de destino (YOSINSKI; CLUNE; BENGIO; LIPSON, 2014).

Em Tan *et al.* (2018), é apresentada uma classificação de modos de transferência de aprendizado baseado em quatro categorias: instância, mapeamento, rede e contradição. Na Figura 18 é apresentado cada um dos modelos disponíveis para essa operação.

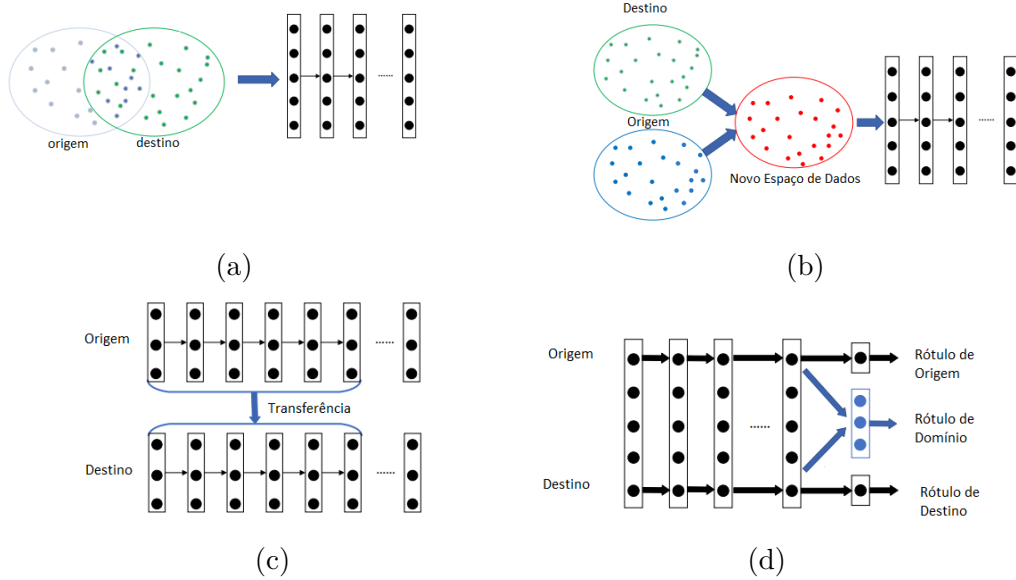


Figura 18 – Representação dos tipos de transferência de aprendizado: (a) baseado em instância, (b) baseado em mapeamento, (c) baseado em rede e (d) baseado em contradição. **Adaptado de:** Tan *et al.* (2018)

A primeira abordagem parte da suposição de que embora haja diferenças entre dois domínios, as instâncias parciais no domínio de origem podem ser utilizadas pelo domínio de destino com pesos apropriados. Essa abordagem explora a estratégia de ajuste de pesos específicos, selecionando instâncias parciais do domínio de origem como suplementos para o treinamento definido no domínio de destino.

A segunda abordagem, denominada transferência de aprendizagem por mapeamento, tem como estratégia a suposição de que embora existam diferenças entre dois domínios de origem, eles podem ser mais semelhantes em um novo espaço de dados. Assim, mapeia o domínio de origem e de destino em um novo espaço de dados. Nesse novo espaço, instâncias de dois domínios são semelhantes e adequadas para o treinamento.

A terceira forma é a transferência de aprendizado baseada em rede. Essa abordagem considera que as redes neurais têm mecanismos de processamento semelhantes ao cérebro humano. Assim, nesse modelo, as camadas frontais da rede podem ser tratadas como um recurso extrator e os recursos extraídos são versáteis. Essa abordagem referencia a reutilização da rede parcial pré-treinada no domínio de origem, incluindo sua estrutura de rede e conexão de parâmetros, transferindo-a para ser uma parte da rede neural profunda que é usada no domínio alvo.

Por fim, a aprendizagem de transferência profunda baseada em contradição baseia-se no pressuposto de que, para uma transferência eficaz, uma boa representação deve ser discriminativa para a tarefa de aprendizagem principal e indiscriminada entre o domínio

de origem e o domínio de destino. Logo, busca encontrar representações transferíveis que são aplicáveis tanto ao domínio de origem quanto ao domínio de destino.

2.9 Comitê de Classificadores

O uso de comitê de classificadores vem demonstrando sucesso em aplicações de aprendizado de máquina nos últimos anos. Devido à capacidade de treinar diferentes modelos e associar as predições, a estratégia possibilita aplicações de análise histológica para obter resultados relevantes (KASSANI; KASSANI; WESOLOWSKI; SCHNEIDER; DETERS, 2019). A união de resultados gerados por dois ou mais algoritmos de classificação, fornece uma tomada de decisão mais efetiva (ROBERTO, 2021). Na Figura 19, é ilustrado um conjunto composto por três classificadores, no qual cada algoritmo apresenta diferentes limiares e o resultado é a composição dos modelos. Dependendo da sua estrutura particular, um comitê pode ser realizado usando duas abordagens principais: heterogênea e homogênea. A primeira abordagem combina diferentes tipos de algoritmos de aprendizado como classificadores individuais, enquanto que a segunda abordagem combina algoritmos de aprendizado do mesmo tipo (NASCIMENTO; CANUTO; COELHO, 2014).

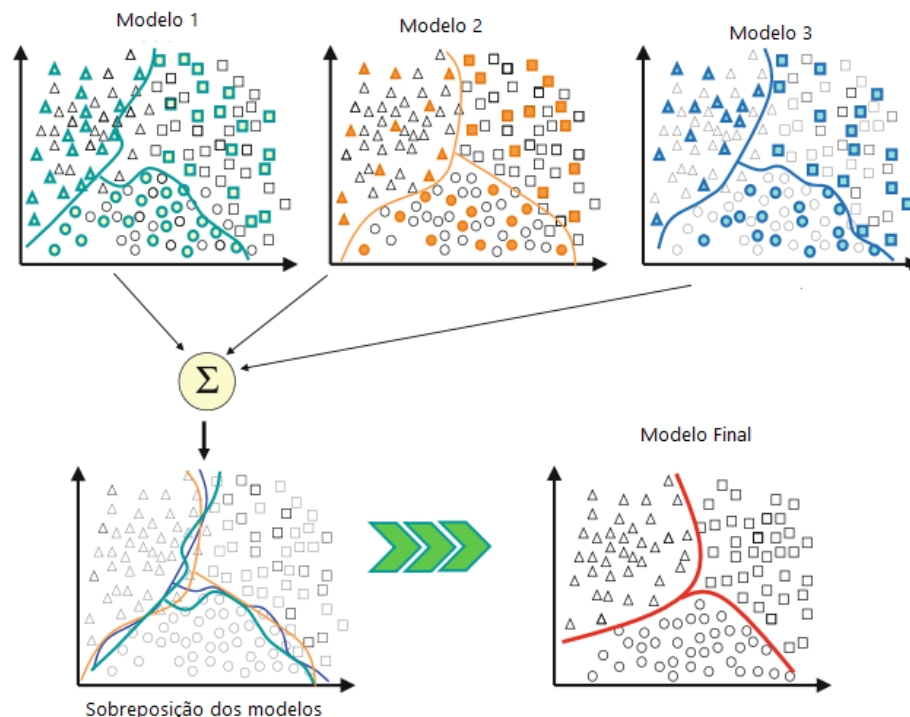


Figura 19 – Comportamento de um conjunto de classificadores composto por três classificadores. **Adaptado de:** (POLIKAR, 2012)

Os autores em Roberto (2021) explicam que existem duas abordagens principais ao utilizar um comitê de classificadores, o *bagging* e o *boosting* (ver Figura 20). O *bagging* consiste em uma classificação paralela de um conjunto de dados usando classificadores

independentes. Já o *boosting* consiste na classificação sequencial do conjunto de dados utilizando os erros dos primeiros modelos treinados para fazer ajustes nos pesos matemáticos dos próximos modelos, visando corrigir classificações erradas (POLIKAR, 2012).

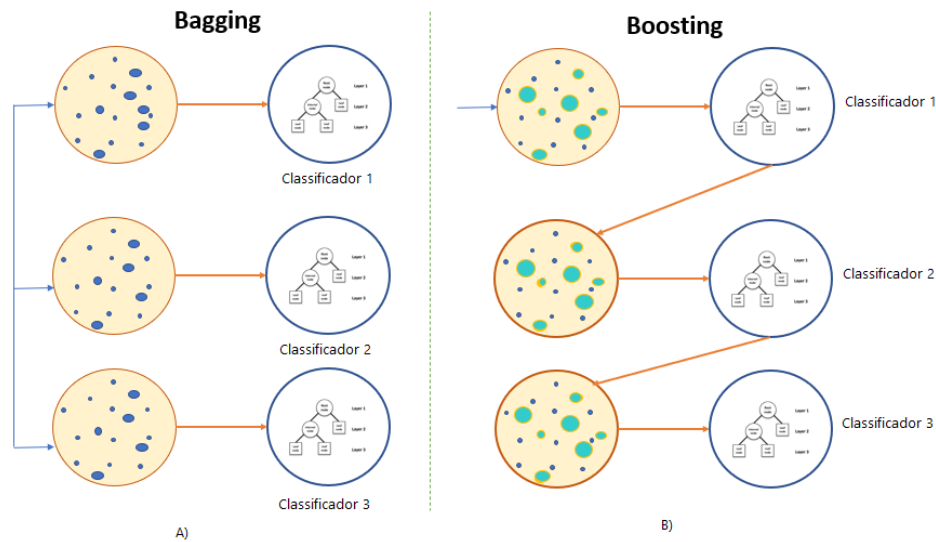


Figura 20 – Abordagens para utilizar comitê de classificadores, (a) *bagging* e (b) *boosting*. Adaptado de: Gaurav (2020)

As técnicas baseadas em *boosting* têm um melhor desempenho em classificações binárias, porém, é mais suscetível a *overfitting* e a ruído, visto que um mesmo conjunto de dados é retreinado várias vezes. Já o *bagging* geralmente tende a produzir melhores modelos comparado a utilização dos mesmos classificadores aplicados isoladamente.

Assim, além do tipo de combinação e da utilização de diversos classificadores distintos ou semelhantes, existem várias abordagens para realizar a combinação dos resultados obtidos pelos diferentes classificadores. Entre as abordagens, destaca-se a votação da maioria, a regra da soma e a soma ponderada. O sistema de votação da maioria realiza a contagem da classe escolhida por cada modelo de classificação. O modelo que tem maior número de votos baseados na rotulação das amostras é escolhido. Essa abordagem apresenta resultados relevantes para classificações binárias (KRAWCZYK; MINKU; GAMA; STEFANOWSKI; WOŹNIAK, 2017). A regra da soma consiste no somatório das probabilidades calculadas pelos classificadores baseados em probabilidades para escolha da classe adequada. Com isso, um resultado errôneo apresentado por determinado modelo pode ser minimizado (SAGI; ROKACH, 2018). A terceira estratégia, denominada soma ponderada, também emprega o somatório das probabilidades de classes, mas inclui valores de pesos para modelos específicos, sendo requerido uma etapa de pré-treino para definir os pesos (NANNI; LUMINI; GHIDONI, 2018).

2.10 Algoritmos Bio-inspirados

Uma outra abordagem utilizada para melhorar os resultados de classificação com CNNs é o uso algoritmos bio-inspirados para otimização de parâmetros das redes. Esses métodos têm recebido grande atenção, nos últimos anos, e buscam compreender e aplicar os padrões encontrados na natureza, com algumas simplificações, em problemas reais. A origem desses algoritmos derivam do estudo de sistemas naturais, e sua eficácia é relacionada à imitação das características mais vantajosas encontradas na natureza. Um exemplo é a seleção do indivíduo mais apto e a capacidade de adaptação ao ambiente, características desenvolvidas ao longo de milhões de anos por meio da seleção natural. (BARBOSA, 2017). As abordagens de algoritmos podem ser divididos em: algoritmos genéticos, algoritmos bioinspirados baseados em inteligência de enxame, algoritmos bioinspirados baseados na ecologia e algoritmos bioinspirados multiobjetivos. Nesse trabalho, as investigações ocorreram baseados no modelo genético.

2.10.1 Algoritmo Genético

O Algoritmo Genético (AG) é um algoritmo inspirado na teoria da evolução introduzido por Charles Darwin. Esse método é baseado na ideia da seleção natural, inspirado em conceitos da genética e cruzamento de indivíduos mais aptos para alcançar a solução ótima do problema a ser tratado (TOSTA *et al.*, 2016). É considerado uma ferramenta promissora na busca por soluções otimizadas em um domínio, inicialmente, desconhecido de maneira altamente eficiente (ABBASGHOLIPOUR; OMID; KEYHANI; MOHTASEBI, 2011).

Assim, essa abordagem, inicialmente, tem uma população de possíveis soluções do problema estudado que evolui ao longo de gerações resultando em descendentes com genes (atributos) ótimos. É importante que no problema analisado haja uma tendência de convergência em direção às melhores soluções a cada geração, uma vez que ao cruzar indivíduos ótimos se espera encontrar uma combinação ainda melhor para os descendentes (MENDES *et al.*, 2021). Contudo, nessa abordagem, é possível que ocorra um direcionamento para uma solução ótima local. Assim, é necessário adicionar aleatoriedade para evitar esse comportamento. Portanto, um AG é composto por alguns conceitos similares ao da biologia como:

- ❑ Gene: Representa um atributo único de um cromossomo;
- ❑ Cromossomo: Representa um conjunto de genes ou atributos que diferenciam os indivíduos;
- ❑ Indivíduo: É a composição de um ou mais cromossomos que representam uma potencial solução para o problema;

- ❑ População: É o conjunto de indivíduos a serem manipulados durante a cada geração do algoritmo;
- ❑ Aptidão: É a medida empregada para avaliar a competência de um indivíduo dentro do espaço de busca;
- ❑ Seleção: É o processo de definição dos melhores indivíduos que irão cruzar para gerar novos indivíduos;
- ❑ Cruzamento: É o processo utilizado para combinar os genes do indivíduos gerando novos descendentes;
- ❑ Mutação: É o processo de modificar indivíduos para introduzir diversidade na população, prevenindo a convergência para ótimos locais;
- ❑ Geração: É o processo de iteração do método para as atualizações.

Na Figura 21 é apresentado um fluxograma de execução do algoritmo AG. A função objetivo avalia a aptidão da população, ou seja, definirá o seu desempenho. Se os critérios de paradas não forem atingidos, esses indivíduos passam por um processo de seleção, cruzamento dos genes, semelhante a uma reprodução, e ainda mutação, formando assim uma nova população. Essa nova população irá passar pelos mesmos processos até que atinga algum critério de parada, que pode ser, o tempo de execução, o número de gerações, a diferença média entre as gerações de indivíduos ou um valor objetivo.

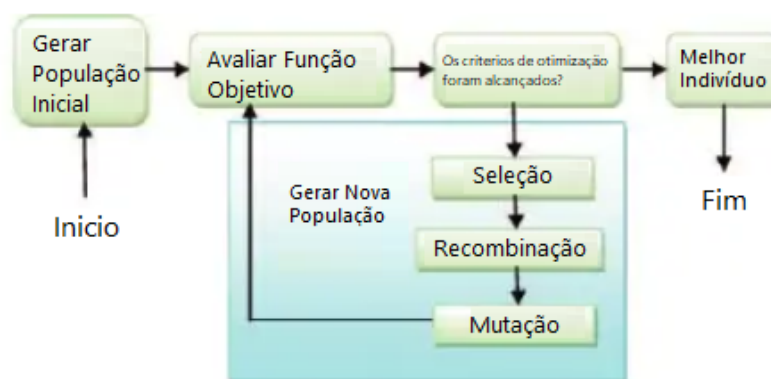


Figura 21 – Fluxograma das etapas de um AG. **Fonte:** (ABBASGHOLIPOUR; OMID; KEYHANI; MOHTASEBI, 2011).

Essa técnica pode ser empregada para definição do número de atributos em um processo de seleção de característica e também pode ser empregado na definição de parâmetros mais adequados no processo de otimização de uma solução como os hiperparâmetros de CNNs (GONÇALVES; SOUZA; FERNANDES, 2022).

2.11 Considerações Finais

Neste capítulo, foram abordados os conceitos teóricos relacionados ao tema do estudo e à metodologia proposta. Esses tópicos englobam as etapas e abordagens utilizadas no trabalho, sendo fundamental compreender as características e os desafios específicos associados a cada tópico abordado para o desenvolvimento das próximas etapas da pesquisa. Destaca-se a importância dos atributos fractais e da representação desses atributos em matrizes de 2 dimensões. Além disso, é importante ressaltar os modelos de CNNs com uso dos sistemas de comitês de classificadores bem como a otimização de seus hiperparâmetros via algoritmos bio-inspirados.

Trabalhos Relacionados: Estado da Arte

Nesta seção, são apresentados os estudos recentes relacionados a detecção de displasia e ao processamento de imagens histológicas utilizando propriedades fractais e comitê de classificadores. Os trabalhos são descritos para embasar o estudo proposto.

3.1 Detecção e Classificação de Displasias

Dentro da área de saúde oral, as displasias são investigadas para detectar as características e anomalias que contribuem de forma significativa no desenvolvimento e detecção dessa patologia.

Dentre soluções apresentadas na literatura, destaca-se o trabalho proposto pelos autores em Krishnan *et al.* (2009). Esse estudo apresenta um método para classificação automatizada de células de tecidos com Fibrose Submucosa Oral (FSO) com exploração de características geométricas. Nesse trabalho, 2.894 células foram, manualmente, identificadas e classificadas por dois especialistas de forma independente. Foi empregado o algoritmo SVM como classificador em que 1.054 células foram usadas no conjunto de treinamento e 1.840 células para o conjunto de teste utilizando o método de validação cruzada. Dentre os dados, 730 células eram de amostras normais e 1.110 células estavam relacionadas a FSO. O método proposto pelos autores, apesar de focar em características geométricas obteve uma acurácia de 88,69%. Os autores concluíram que é possível rastrear uma divergência biológica por meio de técnicas matemáticas e, no futuro, pretendem explorar outras características morfológicas.

Outro estudo apresentado por Krishnan *et al.* (2011) propôs um método para classificação de FSO baseado em informações de textura. Nesse estudo, foram utilizadas 158 imagens, sendo 90 de tecidos normais, 42 de FSO sem displasia e 26 de FSO com displasia. Para obtenção dos descritores de características, os autores usaram a dimensão fractal, padrões binários locais e transformada wavelet. Na etapa de classificação, uma variação do classificador SVM, denominada *one-against-one SVM*, foi empregada para caracterizar as imagens entre saudáveis e doentes. O sistema obteve acurácia de 91,67%, 85,71% e

87,76% na classificação de imagens normais, com FSO sem displasia e FSO com displasia, respectivamente. Em relação aos atributos, os autores verificaram que a combinação de medidas de textura de wavelet e padrões binários locais foram as informações mais relevantes para caracterizar os subtipos das lesões investigadas. Em trabalhos futuros, os autores pretendem investigar o método em outras lesões histopatológicas semelhantes.

Em Baik *et al.* (2014), os autores desenvolveram um método para diferenciar lesões pré-malignas orais, com alto risco de progredir para carcinoma, daquelas de baixo risco. Os atributos extraídos das imagens foram dados morfométricos. Um classificador de núcleos baseado em floresta aleatória foi treinado para discriminar 8.841 núcleos anormais de 5.762 núcleos normais. Os núcleos foram obtidos de 133 amostras de tecidos histológicos da cavidade bucal, sendo que 62 amostras compuseram o grupo de treinamento e 71 imagens definiu o grupo de teste. O algoritmo mostrou uma classificação correta com taxa de 80% para o conjunto de teste. Os autores descreveram que o método proposto serviu como um recurso crucial na implementação de análise de imagens de alta resolução em patologia clínica para identificar lesões que requerem avaliação molecular ou acompanhamento mais frequente.

Os autores em Adel *et al.* (2018b), introduziram uma classificação de FSO usando os algoritmos, denominados *Oriented FAST e Rotated BRIEF* (ORB), *Scale Invariant Feature Transform* (SIFT) e a técnica *Speeded Up Robust Features* (SURF), para extração de características e para etapa de classificação empregaram o SVM. O método proposto ainda apresentou uma etapa de pré processamento convertendo cada imagem de RGB para o formato HSV e uma etapa de redimensionamento ajustando as imagens para 600×600 pixels. A técnica proposta foi treinada e testada em um total de quarenta e seis casos em imagens com uma magnificação de $100\times$. As imagens foram divididas em dois grupos aleatórios, de treinamento e de teste, equivalentes a 32 (70%) imagens e 14 (30%) imagens, respectivamente. A abordagem proposta alcançou 92,8% de precisão em classificação de displasia epitelial oral. Para futuros trabalhos, os autores propuseram aumentar o número de casos baseado em parcerias com outras instituições com objetivo de validar a relevância da abordagem para o tratamento patológico automatizado de displasia bucal.

Em Huttunen *et al.* (2020), os autores investigam um método para classificação automática de displásicas com alta precisão, usando CNNs pré-treinadas. O principal diferencial desse trabalho é uso da microscopia multifotônica na captação das imagens. De acordo com os autores, essa abordagem pode melhorar o desempenho em relação algumas das desvantagens específicas da histopatologia tradicional, explorando vários sinais ópticos endógenos para fornecer biópsias virtuais que refletem a arquitetura e a composição dos tecidos. Além de serem de fácil interpretação por um especialista treinado, podem ser classificadas automaticamente em classes saudáveis ou com lesão. Para classificar as imagens coletadas foi empregado a CNN pré-treinada GoogLeNet com uso de transferência de aprendizado para extrair características e realizar uma classificação binária da ima-

gens. Os resultados obtiveram sensibilidade, especificidade e precisão superiores a 90%. Os autores sugerem que o uso da microscopia multifotônica combinada com um modelo de aprendizado profundo para rastreamento de displasias pode facilitar o diagnóstico e a intervenção preventiva, o que permite abordagens terapêuticas mais adequadas.

Em Silva *et al.* (2022), os autores apresentam um método para quantificação de displasia em imagens histopatológicas da cavidade oral utilizando modelos de aprendizado de máquina. A metodologia inclui as etapas de segmentação de núcleos, pós-processamento, extração de características e classificação. Na etapa de segmentação, foi empregada a rede neural Mask R-CNN. A etapa de pós-processamento empregou operações morfológicas para remover áreas de falsos positivos e negativos. Em seguida, 23 características morfológicas e não morfológicas como área, orientação, solidez e entropia foram computadas e um classificador polinomial foi empregado para distinguir as imagens entre os graus da lesão. Essa abordagem foi aplicada em um conjunto de dados com 296 regiões de interesse da imagem, em que 9.155 núcleos celulares foram identificados e analisados. A etapa de segmentação resultou em acurácias que variaram de 88,92% a 90,35% e a etapa de classificação obteve-se uma área sob a curva ROC de 0,970. Os autores afirmaram que a partir desses valores, a metodologia proposta contribuiu para o estado da arte e pode ser utilizada como ferramenta para auxiliar os patologistas com valores precisos para investigação de lesões displásicas.

Já no trabalho de Neves *et al.* (2023), os autores estudaram uma técnica de aprendizado de máquina baseado na associação de redes neurais convolucionais e representações DeepDream com múltiplos classificadores. Nesse trabalho, foi desenvolvido um estudo para classificar padrões em imagens histológicas, dentre elas, imagens de displasia oral classificadas de forma binária, normal e com displasia. Os autores extraíram diversas características e utilizaram o algoritmo Relief para selecionar as mais relevantes. Para etapa de classificação vários algoritmos foram investigados, porém, o modelo Decorate obteve o resultado mais relevante, alcançando uma acurácia de 97,97% para esse tipo de lesão. Segundo os autores, esse estudo permite expandir o uso de técnicas consolidadas na área de classificação e reconhecimento de padrões na classificação de imagens H&E.

Por fim, no trabalho de Silva *et al.* (2024), os autores publicaram um conjunto de imagens públicas referentes a displasia epitelial oral. O conjunto de dados inclui 456 imagens adquiridas de 30 línguas de camundongos. As imagens foram categorizadas entre os graus de lesão, com estruturas nucleares marcadas manualmente por um especialista treinado e validadas por um patologista. Além disso, os autores também realizaram experimentos para ilustrar o potencial do conjunto de dados proposto em processos para segmentação semântica e de instância. Também foram avaliados o pré-processamento com métodos de normalização de coloração H&E. Na etapa de classificação, as imagens foram classificadas com arquiteturas de CNN e outros algoritmos de aprendizado de máquina. Na etapa de classificação, o resultado mais expressivo foi alcançado com o classificador

Random Forest, com um valor de acurácia de 94,22%. Os autores destacam que, apesar dos resultados encontrados, novos estudos são necessários para o aprimoramento dessas etapas de diagnóstico automatizado. Assim, a disponibilização do banco de imagens permite que outros pesquisadores investiguem outras técnicas de classificação, aprimorando o uso de CAD em aplicações clínicas de imagens de tecido de displasia epitelial oral.

3.2 Geometria Fractal

Um dos estudos iniciais proposto com um método para análise de estruturas multi-fractais para identificar imagens histológicas foi em Vasiljevic *et al.* (2012). Nesse estudo, os autores trabalharam com imagens de câncer primário em metástases ósseas. Essas imagens de tecidos histológicos foram coloridas com H&E. Ao todo, foram analisadas imagens de 1.050 pacientes com casos de câncer de mama, pulmonar e renal. Os atributos multi-fractais foram extraídos a partir dos pontos máximos e mínimos das funções geradas pela DF generalizada e pela análise do espectro multi-fractal. Essa análise consiste em obter diferentes valores de dimensão fractal para as regiões das imagens, ao invés de obter apenas o valor global do atributo. Esses atributos foram classificados por uma árvore de decisão, gerando uma acurácia de 73,38%, 73,71% e 76,22% para as classificações de câncer de pulmão, mama e rim, respectivamente. Apesar do estudo apresentar valores regulares de acurácias, os autores concluíram que a classificação de imagens microscópicas médicas por análise multi-fractal foi promissora e poderia contribuir em casos em que o diagnóstico não era preciso.

Outra propriedade da geometria fractal, a LAC foi aplicada em Neves *et al.* (2014) para diferenciação entre casos de hiperplasia, neoplasia e tecidos saudáveis em imagens histológicas de próstata. Os autores utilizam o algoritmo *glinding-box* para extrair valores locais e globais de LAC. Esses valores foram usados para elaborar os vetores de características, os quais foram dados como entrada para uma árvore de decisão e um SVM com kernel polinomial. O melhor resultado foi de 90,63% na comparação entre tecido normal e tecido neoplásico utilizando o classificador polinomial. Dentre os desafios em aberto, os autores afirmaram que a associação da LAC com os demais descritores para composição do vetor de características poderá contribuir com melhores resultados.

Em Lennon *et al.* (2015), os autores realizaram uma análise quantitativa de valores de DF e LAC em imagens de câncer pulmonar e de tecido saudável. Para a análise quantitativa, as métricas de DF e LAC foram extraídas por meio da abordagem *box-counting*. Em uma das principais observações dos autores, o valor da DF era maior para as imagens com casos de câncer quando comparados às amostras saudáveis. Enquanto que o valor de LAC apresenta o comportamento inverso. Os autores, ao investigar na literatura, relataram que a medida da DF não era capaz de promover uma boa separação entre subtipos de câncer, assim combinar com outros atributos melhorou a diferenciação

dos tipos de câncer pulmonar. Dentre os desafios encontrados na pesquisa, os autores destacaram a necessidade de padronizar as técnicas propostas na literatura para cálculo desses atributos, uma vez que há diversas implementações para cálculo DF e cada um delas apresenta resultados discordantes, o que pode comprometer a análise de uma imagem.

Além das métricas DF e LAC, a PERC é um outro descritor fractal que pode ser utilizado para extração de características de imagens. O trabalho proposto por Roberto *et al.* (2017) investigou os atributos globais de percolação de imagens histológicas de diferentes classes de linfomas. Os atributos locais foram obtidos por meio da abordagem *gliding-box*. Dentre os atributos extraídos, a técnica proposta obteve como métricas o número de aglomerados, ocorrência de percolação e ocupação do maior aglomerado. Para etapa de classificação, diversos classificadores foram investigados: K^* , *Logistic*, *Multilayer Perceptron*, *RBFNetwork*, *DECORATE*, *Logit Boost*, *MultiClass Classifier*, *Random Committee*, *Rotation Forest* e *Random Forest*. Foi realizada uma amostragem baseada em *10-fold cross-validation* com 90% dos dados correspondendo ao grupo de treinamento e 10% restante a validação. O melhor resultado foi obtido por um classificador baseado em *Random Forest* em imagens segmentadas, com um valor de acurácia médio de 92%. Por ser ainda pouco explorada, a PERC pode ser aplicada em outros contextos e a associação da PERC com atributos tradicionais da geometria fractal ou em conjunto com outros atributos pode melhorar ainda mais um sistema CAD na classificação de lesões em imagens histológicas.

Ainda no contexto de atributos fractais, o trabalho de Pereira *et al.* (2022) trouxe contribuições à literatura ao propor um sistema para classificação de imagens histológicas de linfoma. Esse trabalho propôs um método baseado na associação de características extraídas por técnicas fractais, regularização e classificador polinomial. Os vetores de características foram classificados aplicando a técnica de validação cruzada com *10 fold*. A abordagem proposta alcançou resultados significativos para imagens de lesões de linfomas não Hodgking, alcançando valores em torno de 99% de acurácia. De acordo com os autores, esses resultados foram considerados relevantes em relação aos estudos na literatura. A associação do polinômio de Hermite e regularização pode contribuir para a detecção das lesões, apoiando especialistas em práticas clínicas.

O trabalho de Longo *et al.* (2023) propôs um sistema para identificar as melhores combinações de descritores tradicionais e descritores de aprendizagem profunda para classificar imagens histológicas. Os descritores tradicionais incluem técnicas fractais multiescala e multidimensionais, como DF, LAC e PERC, enquanto as características aprendidas foram extraídas de várias redes neurais convolucionais. Os autores investigaram diferentes associações entre esses descritores, aplicando-as em um conjunto diversificado de classificadores. O método foi avaliado em diferentes imagens histológicas como câncer de mama, câncer colorretal, displasia oral e tecido hepático. O método proporcionou taxas de acurácia em torno de 94,83% e 100%. Os autores concluem que esses resultados contri-

buem para avançar o desenvolvimento de sistemas de auxílio ao diagnóstico para imagens histológicas.

Já o trabalho de Roberto *et al.* (2023), os autores investigaram o uso de atributos de PERC em imagens digitais de raios X para detecção de Covid-19. Nesse estudo, foi proposto um classificação baseado na extração de características usando os descritores de PERC locais e globais. O método foi aplicado em dois conjuntos de dados: um contendo 2.002 amostras segmentadas divididas em duas classes (Covid-19 e saudável); e outro contendo 1.125 amostras não segmentadas divididas em três classes (Covid-19, saudável e pneumonia). Diversos classificadores foram investigados, empregando o método de validação cruzada com *10 folds*. Dentre os resultados, destaca-se o uso do classificador polinomial de Hermite. Esse classificador forneceu os melhores resultados, alcançando 97,8% de acurácia. Assim, de acordo com os autores, esses resultados obtidos pela associação de características fractais e polinômio de Hermite podem contribuir para a detecção das lesões, apoiando especialistas em práticas clínicas.

3.3 Algoritmos Bioinspirados para Otimização de Modelos de CNNs

Conforme explicitado no capítulo anterior, o desenvolvimento de uma CNN envolve a otimização de vários parâmetros, bem como a escolha de sua arquitetura. Neste contexto, diversos trabalhos têm sido desenvolvido focados na otimização de modelos de CNNs.

Em Faria, Julia, Fernandes, Nascimento e Julia (2023), os autores propuseram a otimização da arquitetura da uma CNN no contexto de identificação de informações relevantes sobre eventos de jogos em tempo real. Esses eventos normalmente são associados a ações realizadas pelos jogadores. A rede investigada para o contexto proposto foi a CNN baseada em redes residuais. Os autores investigaram o uso de um AG e uma otimização Bayesiana para tratar os hiper-parâmetros de uma rede CNN. Dentre os hiper-parâmetros investigados, foram avaliados o número e o tamanho dos filtros das camada convolucionais e a camada de pooling. O critério utilizado para definição do indivíduo mais apto foi uma relação entre o tamanho da rede e a acurácia do modelo. O modelo otimizado pelo AG alcançou 99,35% de acurácia com um tamanho menor que o modelo otimização pela otimização Bayesiana. Os autores concluíram que o método com GA foi superior ao modelo Bayesiano no conjunto de dados analisados.

Em relação ao domínio de imagens médicas, o estudo de Gonçalves, Souza e Fernandes (2022) abordou a detecção de câncer de mama em imagens infravermelhas. Nesse estudo, os autores exploraram duas técnicas de otimização bioinspiradas, o AG e otimização por enxame de partículas (PSO), com o objetivo de encontrar hiper-parâmetros ótimos e as melhores arquiteturas para as camadas totalmente conectadas das CNNs. A base de dados utilizada nesse estudo foi a base pública DMR-IR, composta por 864 imagens saudáveis

e 193 imagens não saudáveis. Para utilizar as CNNs, foi necessária uma etapa de pré-processamento a fim redimensionar as imagens para a camada de entrada das redes. A estratégia explorou apenas a parte das camadas totalmente conectadas, os parâmetros como taxa de aprendizagem, número de neurônios de cada camada totalmente conectada, presença da camada de *dropout* e a taxa de *dropout*. De acordo com autores, a otimização de GA e PSO melhoraram a maioria dos resultados e destacaram que o GA superou o PSO nos experimentos realizados proporcionando um F-score de 0,92 com a VGG-16 e 0,90 com ResNet-50.

O estudo de Najaran (2023) desenvolveu e avaliou um método para detecção automática de casos de Covid-19 por meio de imagens de raios X. Os autores desenvolveram um algoritmo baseado em CNN e um algoritmo evolutivo. O processo de otimização da arquitetura CNNs consistiu em encontrar o número de camadas e a combinação entre elas adequada para o problema. Um modelo de evolução foi proposto para a otimização da estrutura das CNNs. O processo de otimização consistiu dois *threads* que realizaram pesquisas simultaneamente para otimizar a arquitetura. Um *thread* foi aplicado para a pesquisa da ordenação combinatória de operadores convolucionais e de um *pooling*. O segundo *thread* explorou o cenário dos parâmetros numéricos. O processo co-evolutivo proposto variou entre os dois cenários a cada 100 iterações. Assim, nas primeiras 100 iterações, o processo de otimização foi realizado para otimizar a arquitetura das CNNs enquanto os parâmetros dos operadores são fixos. Após 100 iterações, o esqueleto (o gráfico) das soluções foi fixo e a otimização foi realizada para otimizar os parâmetros numéricos. O algoritmo foi aplicado até que o número máximo de iterações. Esse método foi aplicado diversos tipos diferentes de CNNs e obteve resultados significantes com acurácia de 98.25% para detecção dos casos de Covid-19.

3.4 Comitê de Classificadores

Nos últimos anos, métodos baseados em CNN têm apresentado um desempenho promissor para classificação em imagens histológicas. Contudo, o desempenho desses modelos estão sujeitos a diversos problemas, como o desbalanceamento de amostras de cada classe investigada, assim como, a maldição da dimensionalidade podem comprometer uma classificação (SAGI; ROKACH, 2018). Nesse sentido, estudos recentes têm explorado a modelos de comitês de CNNs para aprimorar o desempenho das abordagens propostas. Nesta subseção, são descritos os estudo direcionados a essa estratégia para imagens médicas.

Em Roy, Banik, Bhattacharjee e Nasipuri (2019), os autores propuseram um método baseado em *patch* extraído de uma CNN para classificação automática de imagens histopatológicas de mama, usando a base dados conhecida como ICIAR 2018. O sistema de classificação proposto funciona em dois modos diferentes: um *patch* baseado em uma decisão (do inglês, *one patch in one decision* - OPOD) e todos os *patches* em uma tomada

de decisão, (do inglês, *all patches in one decision* - APOD). O método OPOD avalia a predição dos rótulos das classes. Se esse rótulo de classe for o mesmo para todos os *patches* extraídos e representar o rótulo de classe dessa imagem, a saída era considerada como classificação correta. No método APOD, o rótulo de classe de cada *patch* foi extraído com os mesmos procedimentos do OPOD e um esquema de votação por maioria permitiu definir o rótulo de classe da imagem. Os resultados experimentais mostraram que o modo OPOD proposto alcançou uma precisão de classificação de 77,4% para quatro rótulos e 84,7% para duas classes de lesões histopatológicas. Enquanto que o APOD alcançou precisão de classificação de imagem de 90% para quatro classes e 92,5% para classificação de duas classes. Essa abordagem mostra que o uso de um conjunto de classificadores, como o modo APOD, pode melhorar o desempenho na obtenção de resultados.

No trabalho de Kassani, Kassani, Wesolowski, Schneider e Deters (2019), foi utilizado um modelo de comitê composto por CNNs pré-treinadas, formadas pelas seguintes arquiteturas: VGG19, MobileNet e DenseNet. No método proposto, as camadas totalmente conectadas finais de cada CNN foram combinadas para produzir um vetor de características. Essa abordagem permitiu explorar as vantagens das diferentes arquiteturas para melhorar a precisão do modelo. Por fim, os atributos extraídos foram utilizados em um classificador de multicamada *perceptron*. Esse modelo foi validado com quatro diferentes bancos de imagens de câncer de mama disponíveis na literatura: ICIAR, BreakHis, PatchCamelyon e Bioimaging. O método obteve melhores resultados ao comparado com classificadores separados, obtendo precisão de 98,13%, 95,00%, 94,64% e 83,10% para os conjuntos de dados de BreakHis, ICIAR, PatchCamelyon e Bioimaging, respectivamente. Os autores concluíram que utilizar um conjunto multi-modelo baseado em aprendizado profundo melhora o desempenho em relação as técnicas separadas.

O trabalho de Roberto (2021) propôs um modelo de conjunto de classificadores baseado em atributos fractais e aprendizado profundo. Esse modelo consistiu em combinar a classificação de duas CNNs, a *F-CNN* e a *O-CNN*, em imagens de diferentes classes de tecido da histopatologias. Na extração das características, foi aplicada geometria fractal para obter 300 atributos fractais de diferentes conjuntos de dados histológicos. Esses atributos foram reorganizados em um processo conhecido como *reshape*. Nesse processo, esses atributos são transformados em uma matriz a fim de compor uma imagem RGB. Essa imagem gerada foi utilizada como entrada para classificação pela *F-CNN* e a *O-CNN*. O comitê de classificadores foi obtido da associação da *F-CNN* e a *O-CNN* usando a regra da soma ponderada. Nos resultados, foram obtidas acurácias entre os intervalos de 88,45% até 99,77%. Além disso, o modelo foi capaz de classificar imagens de conjuntos de dados com classes desbalanceadas, o qual é um dos grandes problemas para um modelo de aprendizado de máquina.

3.5 Considerações finais

Esse capítulo apresentou várias abordagens recentes na literatura para a classificação de imagens histológicas, com foco principalmente na detecção de displasias orais. Os pesquisadores têm buscado maneiras de aprimorar a detecção e classificação dessas lesões, visando melhorar o diagnóstico e o tratamento. Dentre as técnicas utilizadas, a extração de atributos fractais, como a dimensão fractal, lacunaridade e percolação, tem se mostrado promissora na análise de imagens histológicas. Esses descritores têm sido associados a vários algoritmos de classificação, desde árvores de decisão até redes neurais convolucionais. O uso de algoritmos bioinspirados, como algoritmos genéticos, tem demonstrado uma relevante estratégia para a otimização de parâmetros em arquiteturas de CNNs podendo contribuir com os modelos de classificação. Ainda, o uso de diferentes CNNs em um comitê de classificadores tem se mostrado importante na melhoria da precisão na classificação de imagens histopatológicas. O uso dessas técnicas, no contexto de displasia, ainda não foi explorado para avaliação dos graus do avanço da doença. Investigar novas abordagens para estudo de modelos mais robustos e precisos são relevantes para área médica.

Metodologia

Nesta seção é apresentado o banco de imagens empregado na investigação dos modelos propostos, assim como, os algoritmos computacionais propostos para obtenção dos descritores e classificação das lesões histológicas.

4.1 Introdução

Este capítulo traz uma descrição das etapas exploradas no desenvolvimento da metodologia do trabalho. Os estágios da abordagem são: i) a extração de características fractais, reorganização da informação através da representação de características; ii) otimização de modelos de redes CNNs; iii) comitês de classificadores baseados em modelos de redes CNNs. As principais etapas são representadas na Figura 22.

Nesta etapa, o algoritmo proposto foi desenvolvido com uso da linguagem Matlab, empregando a extensão *Deep Network Designer*, e partes de códigos em python. A avaliação estatística foi realizada por meio do software Jamovi. Todos os experimentos foi executados em um notebook Acer Predator Helios 300 com 32GB de memória RAM e processador de 2,6GHz Intel(R) Core(TM) i7-9750H.

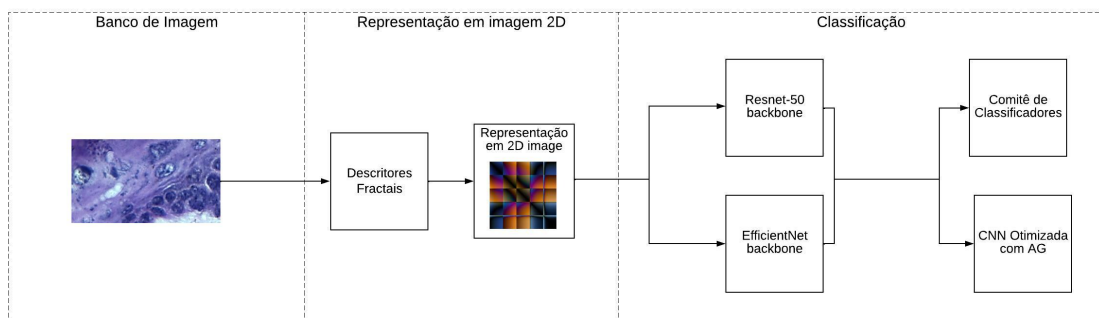


Figura 22 – Visão geral da metodologia proposta baseada em representação de atributos fractais para classificação de displasia de tecidos da cavidade oral. **Fonte:** Elaborado pelo autor.

4.2 Banco de Imagens

O banco é composto por imagens microscópicas de lesões da língua de camundongos. Essas imagens foram obtidas por meio de experimento in vivo devidamente aprovado pelo Comitê de Ética na utilização de animais sob número de protocolo 038/09 e realizado no período de 2009 a 2010. Essas imagens já foram empregadas nos trabalhos apresentados por Sant’Ana *et al.* (2011), Faria *et al.* (2011), Silva *et al.* (2019), .

Na obtenção das imagens, foi empregado o protocolo de indução de lesões em língua de camundongos. Nesse caso, foram usados 30 animais pesando entre 20 e 23 gramas, de idade de 6 a 8 semanas e distribuídos de forma aleatória em dois subgrupos denominados de semana 16 e semana 32. Todos os camundongos foram colocados em gaiolas de plástico contendo serragem e mantidos à temperatura de 22° C, ciclo claro/escuro de 12h, no centro de bioterismo da Universidade Federal de Uberlândia. O protocolo obedeceu os critérios estabelecidos pelos autores em Tang, Knudsen, Bemis, Tickoo e Gudas (2004), como segue: o carcinógeno 4NQO foi previamente diluído em propilenoglicol na concentração de 5 mg/ml e depois em água filtrada até atingir a concentração final de 100 ug/ml. Após o sacrifício de ambos os grupos, as línguas foram removidas e colocadas em solução de formalina tamponada a 4% por 24 horas. Depois desse período, as línguas foram processadas e incluídas em parafina para confecção dos blocos, recortadas e coradas por hematoxilina e eosina (H&E) para estudo histopatológico. Com ajuda de um patologista, as lâminas foram classificadas em tecido saudável, displasias leve, displasia moderada e displasia grave.

As imagens histológicas digitais foram obtidas com o uso do microscópio óptico Leica DM500, em magnificação de 400x, e armazenadas em arquivos em formato TIFF, utilizando o modelo de cores RGB, com quantificação em 8 bytes e resolução de 2048×1536 pixels. Ao todo, 43 lâminas foram digitalizadas e as imagens foram recortadas em regiões de interesse (do inglês, *region of interest* - ROIs) com tamanho de 450×250 pixels, sendo 74 ROIs para cada classe de tecido, totalizando 296 ROIs. Na Figura 23, é apresentado um exemplo de cada tipo de tecido (displasia leve, moderada e severa e tecido saudável) disponível no banco de imagens.

4.3 Atributos Fractais

Conforme mencionado na Seção 2.4, uma das principais etapas no desenvolvimento de uma sistema CAD com imagens histológicas é a utilização de atributos fractais globais e locais. Portanto, para construção dos vetores foi empregado o método *glinding box* para obtenção da matriz, de probabilidade. Com essa matriz foi possível obter os atributos de DF, LAC e PERC para geração dos vetores de características.

Durante o processo de obtenção das matrizes, diversas distâncias podem ser utilizadas.

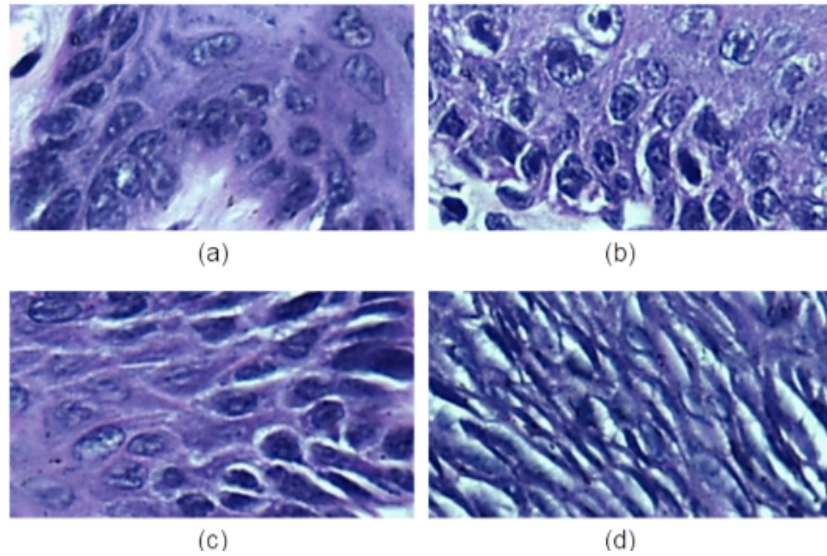


Figura 23 – Exemplos de tecidos histológicos da cavidade bucal: (a) saudável, (b) displasia leve, (c) displasia moderada e (d) displasia severa. **Fonte** : (SILVA *et al.*, 2019)

Neste trabalho, foram empregadas a distância de Minkowski, apresentada na Equação 2, Euclidiana, dada pela Equação 3, com $p = 1$, e Manhattan, descrita na Equação 3, com $p = 2$. Assim, como na distância de Minkowski, o cálculo das demais distâncias foram investigadas a partir da diferença entre o pixel central f_{ci} e os demais pixels f_i em uma região de tamanho $L \times L$. Com as matrizes de probabilidade calculadas, os atributos locais e globais de DF, LAC e PERC, para cada distância, foram obtidos, conforme descrito na Seção 2.4. Esses dados foram armazenados em um vetor com 100 atributos. A Figura 24 apresenta um esquema visual da abordagem proposta.

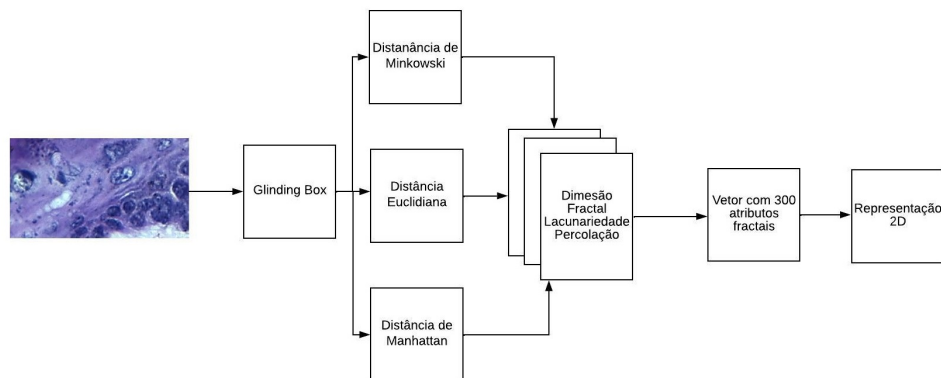


Figura 24 – Visão detalhada da metodologia proposta baseada em representação de atributos fractais. **Fonte**: Elaborado pelo autor.

4.4 *Reshapes* para Representação dos Descritores

Cada característica fractal pode ser obtida de forma global para toda a imagem através da análise de um gráfico em função do tamanho da caixa e de forma local, isto é, para cada tamanho de caixa empregado na construção da matriz de probabilidade. Como a quantidade de caixas depende do tamanho da imagem, conforme Equação 7, os atributos fractais geram um vetor de características. Desta forma, foi utilizado métodos de representação dos atributos 1D em uma matriz 2D.

Diversas abordagens estão disponíveis para a construção de modelos de representação a partir de informações discretas. As representações GR, *Markov Transition Field* (MTF) e *Self-Similarity Matrix* (SSM) são abordagens empregadas para esse propósito. O modelo MTF pode trazer os padrões estatísticos inerentes aos dados. O modelo SSM traz as similaridades entre todos os pares dos elementos dentro de um conjunto de dados, o que permite analisar a redundância ou padrões que se repetem. Além disso, o GR já foi abordado no campo de classificação de lesões histológicas no trabalho de Roberto, Neves, Lumini, Martins e Nascimento (2024), com resultados promissores entre sua comparação e outras abordagens de representação. Então, foi empregado os modelos GR, SSM e MTF para representação das informações em que cada tipo de distância representa um canal de cor RGB. A Figura 25 apresenta o método empregado em que o vetor de atributos calculado pela distância de Minkowski foi reconstruído como canal vermelho (R), o vetor obtido através da distância euclidiana foi definido como canal verde (G), e por fim, o canal azul (B) recebeu o vetor com atributos calculados a partir da distância de Manhattan.

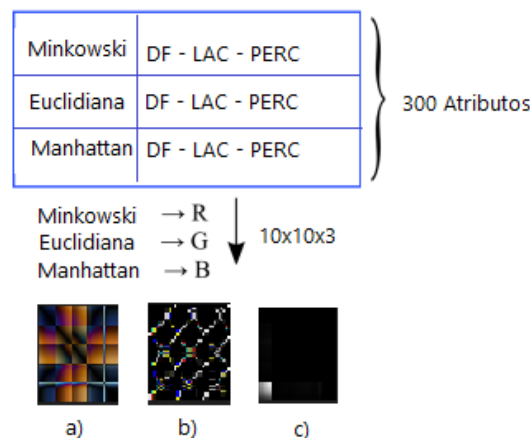


Figura 25 – Exemplo de metodologia proposta ilustrando a separando em cada canal de cor RGB: (a) Representação pelo GR (b) Representação pelo MTF e (c) Representação pela SSM. **Fonte:** Elaborado pelo autor.

4.5 Redes Neurais Convolucionais

Os modelos CNN estão entre as ferramentas mais poderosas para análise de imagens digitais. Neste trabalho, os modelos ResNet-50 (HE; ZHANG; REN; SUN, 2015) e EfficientNet (TAN; LE, 2019) foram aplicados para classificar as imagens de representação devido ao bom desempenho desses modelos na área de classificação de imagens histológicas. Para utilizar ambos os conjuntos de dados nos modelos CNNs, as imagens de entrada foram redimensionadas para 224×224 pixels para ajustá-las a ambos os modelos. Então, próximo passo é estabelecer os conjuntos de dados para etapa de treinamento, validação e teste. O grupo de treinamento é utilizado pela rede durante o aprendizado e a validação é utilizado para validar esse treinamento. Na etapa de treinamento de cada rede, as imagens de representação obtidas após remodelar as características fractais locais foram utilizadas com uma técnica de validação cruzada de 5 *folds*, assim, 20% das imagens foram aleatoriamente selecionadas para compor o grupo de teste e dentre os outros 80% dos dados, 70% ficaram para o grupo de treinamento e outros 30% para a validação. Usou-se ainda 10 épocas e 5 iterações por época com taxa de aprendizado inicial de 0,0001 com a função *piecewise* que reduz a taxa de aprendizado a cada 2 épocas.

Com o objetivo de reduzir o tempo de treinamento da CNN e obter resultados promissores com um menor número de épocas, foi adotada a estratégia de transferência de aprendizagem para diminuir o tempo de avaliação de um modelo sem predefinir os pesos. O modelo de aprendizagem por transferência empregado neste trabalho, foi a transferência baseada em rede, portanto, a última camada totalmente conectada e a camada de classificação para ambos os modelos CNN foram adaptadas com base nas classes do conjuntos de dados.

4.6 Otimização de parâmetros das CNNs

Os autores em Santos (2019) explicam que, com o objetivo de obter melhores resultados, novos modelos apresentam a tendência de construir CNNs cada vez mais profundas e complexas. Entretanto, tais avanços não necessariamente têm feito com que as CNNs sejam mais eficientes em relação ao número de parâmetros, ao custo computacional e ao tempo de inferência. Assim, diversas técnicas de otimização podem ser empregadas para aumentar a eficiência do modelo. Baseado nos resultados obtidos pelo trabalho de Gonçalves, Souza e Fernandes (2022), foi empregado um AG para otimizar a estrutura totalmente conectada das redes, assim foi investigado duas composições, com 2 ou 3 camada totalmente conectadas combinada com uma camada de *dropout* após cada uma das camadas totalmente conectadas. Assim um indivíduo era composto pelos seguintes genes:

□ N1 = Números de neurônios da primeira camada totalmente conectada ([3 10],

inteiro);

- N2 = Números de neurônios da segunda camada totalmente conectada ([3 10], inteiro);
- Dr = Taxa de dropout ([0 1], decimal);
- Peso = Tipos de inicialização de pesos ([1 4], inteiro);
- Bias = Técnica de inicialização de bias ([1 3], inteiro);
- Ta = Taxa de Aprendizado inicial do treinamento ([1,6], inteiro);

Os parâmetro N1 e N2 foram variados no intervalo fechado [3, 10]. O número de neurônios é calculado para cada camada como 2^N . O parâmetro Peso também variou no intervalo fechado [1, 4], em que cada número representa o tipo de inicialização: 1 - *glorot*, 2 - *he*, 3 - *orthogonal* e 4 - *narrow-normal*. Já o parâmetro Bias variou no intervalo fechado [1, 3], em que cada número representa o tipo de inicialização: 1 - *zeros*, 2 - *ones*, 3 - *narrow-normal*. Por fim, o parâmetro Ta variou no intervalo [1,6]. A taxa de aprendizado inicial para cada rede é definida como $10^{(-Ta)}$. O algoritmo foi executado com uma população inicial aleatória composta por 10 indivíduos e foi executado com critério de parada de 10 gerações, assim ao fim da décima geração, o melhor indivíduo foi considerado a rede com parâmetros otimizados.

4.7 Comitê de Classificadores

A utilização de um modelo de comitê está entre as abordagens de maior sucesso em aplicações de aprendizado de máquina nos últimos anos, devido à sua capacidade de treinar diferentes modelos e combinar suas previsões, com resultados relevantes obtidos em histopatologia (KASSANI; KASSANI; WESOLOWSKI; SCHNEIDER; DETERS, 2019). Nesse sentido, para a etapa de classificação, as combinações de diferentes arquiteturas podem melhorar significativamente os resultados. Dentre as diversas técnicas de Comitê, a regra da soma simples é uma técnica de combinação amplamente utilizada em problemas de classificação com modelos CNN. Nesta estratégia, a decisão final do modelo foi obtida calculando a média dos resultados dos modelos individuais (ZHOU, 2012). Assim foi empregado um comitê baseado na regra soma, Equação 16, com pesos (W_i) iguais utilizando as probabilidades de ocorrência (M_i) de cada classe obtidas da saída da camada *Softmax* de cada modelo (n).

$$\sum_{i=1}^n M_i \times W_i \quad (16)$$

Para a avaliação da abordagem proposta, foram investigadas os seguintes comitês:

- modelo 1 - imagens originais e representação GR;

- ❑ modelo 2 - imagens originais e representação MTF;
- ❑ modelo 3 - imagens originais e representação SSM;
- ❑ modelo 4 - representações GR e SSM;
- ❑ modelo 5 - representações GR e MTF;
- ❑ modelo 6 - representações GR, MTF e SSM;
- ❑ modelo 7 - todas as representações.

4.8 Critérios de Avaliação

A avaliação dos métodos propostos foi realizada comparando os resultados obtidos com a classificação padrão-ouro realizada pelo especialista, na qual foram obtidas as métricas de acurácia (A_{CC}), sensibilidade ($Sens$), especificidade (Esp) e medida F1 ($F1score$). A A_{CC} foi utilizada para medir o número de imagens classificadas corretamente de acordo com o padrão ouro, Equação 19, em que verdadeiro positivo (VP) e verdadeiro negativo (VN) representam as imagens classificadas corretamente e falso positivo (FP) e falso negativo (FN) representam as imagens classificadas erroneamente. A especificidade e sensibilidade foram utilizadas para entender a capacidade do método de detectar com sucesso resultados classificados como positivos e negativos. Além disso, a métrica F1 ($F1score$) foi outra métrica importante considerada para avaliar os resultados principais, Equação 20. Por fim, os resultados obtidos foram comparados à trabalhos correspondentes na literatura com o objetivo de verificar o quão eficiente é o método proposto.

$$A_{CC} = \frac{VP + VN}{VP + VN + FP + FN}. \quad (17)$$

$$Esp = \frac{VN}{FP + VN}. \quad (18)$$

$$Sens = \frac{VP}{VP + FN}. \quad (19)$$

$$F1score = 2 \cdot \frac{\frac{TP}{TP+FP} \times \frac{TP}{TP+FN}}{\frac{TP}{TP+FP} + \frac{TP}{TP+FN}}. \quad (20)$$

Experimentos e Análise dos Resultados

Os experimentos para avaliação das abordagens propostas para classificação de lesões da cavidade oral em imagens histológicas são apresentados e discutidos neste capítulo. Para isso, foram realizadas experimentos de classificação das lesões com abordagem denominada referência (*baseline*) e com otimização dos modelos. Também foram investigados a associação das redes convolucionais para a construção de comitês homogêneos na classificação das lesões. Para todos os resultados apresentados, as imagens Originais (O) representam as imagens originais sem nenhum tipo de representação, enquanto que as imagens GR, MTF e SSM são aquelas provenientes da representação de atributos fractais.

5.1 Avaliação dos Modelos de Redes Neurais Convolucionais

Nessa seção, são apresentados os resultados dos experimentos com os modelos tradicionais de arquiteturas das redes convolucionais propostas. A Tabela 2 apresenta os valores da média da acurácia e desvio padrão (σ) do grupo de teste com os *folds* investigados. Os melhores resultados médios obtidos foram destacados em negrito para cada uma das arquiteturas. O tempo médio de treinamento para a classificação foi aproximadamente 40 minutos para arquitetura EfficientNet e 60 minutos para o modelo ResNet-50. Nos resultados da Tabela 2 é possível notar que a utilização da estrutura da ResNet-50 proporcionou valores de acurácia superiores quando comparados ao modelo EfficientNet, alcançando 96,9% com a representação GR dos descritores fractais. Apesar das outras representações fornecerem resultados inferiores ao desempenho quando comparado com a representação GR, é possível notar que o modelo EfficientNet proporcionou resultados superiores para a representação MTF numa comparação com o modelo ResNet. Contudo, mesmo esses valores foram inferiores aos obtidos com a representação GR.

Com objetivo de analisar os detalhes do desempenho dos modelos de CNNs, nas Fi-

Tabela 2 – Resultados da métrica acurácia para as diversas abordagens de entrada para classificação com as redes convolucionais EfficientNet e ResNet-50.

EfficientNet						
	<i>Fold 1</i>	<i>Fold 2</i>	<i>Fold 3</i>	<i>Fold 4</i>	<i>Fold 5</i>	Média (σ)
O	87,0	89,6	92,5	89,2	92,5	90,2 $\pm$ 2,3
MTF	70,6	69,8	74,1	74,2	75,8	72,8 $\pm$ 2,8
SSM	73,2	70,2	79,1	77,5	78,3	75,8 $\pm$ 3,6
GR	97,4	92,2	99,2	96,6	97,5	96,6 $\pm$ 2,6

ResNet-50						
	<i>Fold 1</i>	<i>Fold 2</i>	<i>Fold 3</i>	<i>Fold 4</i>	<i>Fold 5</i>	Média (σ)
O	88,8	91,3	92,5	91,6	90,0	90,9 $\pm$ 1,5
MTF	70,6	69,8	69,1	73,3	76,6	71,9 $\pm$ 3,1
SSM	77,5	77,5	75,8	73,3	79,2	76,7 $\pm$ 2,2
GR	97,4	93,9	97,5	98,3	97,5	96,9 $\pm$ 1,7

guras 26 e 27, são apresentados os gráficos da relação acurácia e número de épocas do primeiro *fold* de treinamento para a representação GR, pois foi o método que mostra os resultados mais relevantes. A escolha do primeiro *fold* foi aleatória, apenas com objetivo de avaliar o comportamento dos modelos durante o treinamento e validação do conjunto de dados.

O treinamento foi realizado com a técnica de validação cruzada com $k=5$, onde 80% das imagens foram selecionadas para o treinamento e 20% para teste dos modelos. No conjunto de treinamento, os dados foram divididos em um conjunto 70% para o treinamento e os outros 30% compuseram o grupo de validação. A partir dos resultados, observa-se que para ambas as estruturas de rede, ResNet-50 e EfficientNet, o grupo de treinamento alcançou 100% de acurácia após a 6^a época. Contudo, nota-se que a ResNet-50 após a época 6^a, houve uma oscilação e retornou para o 100% somente próximo a 10^a época.

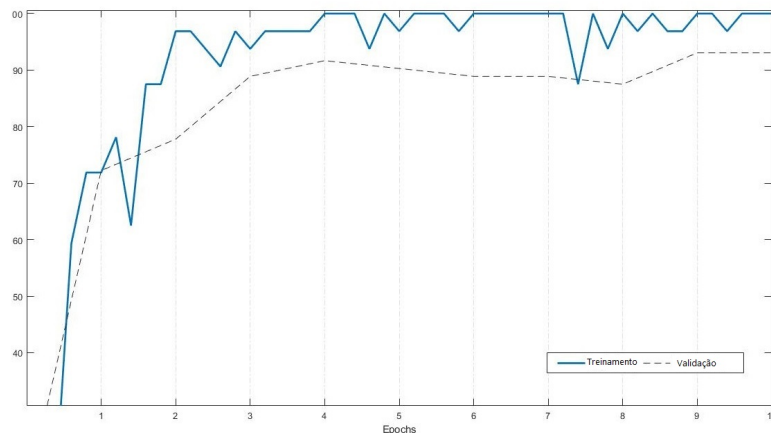


Figura 26 – Avaliação do *fold 1* com representação GR durante o treinamento e validação da rede ResNet-50 *baseline*.

Fonte: Elaborado pelo autor.

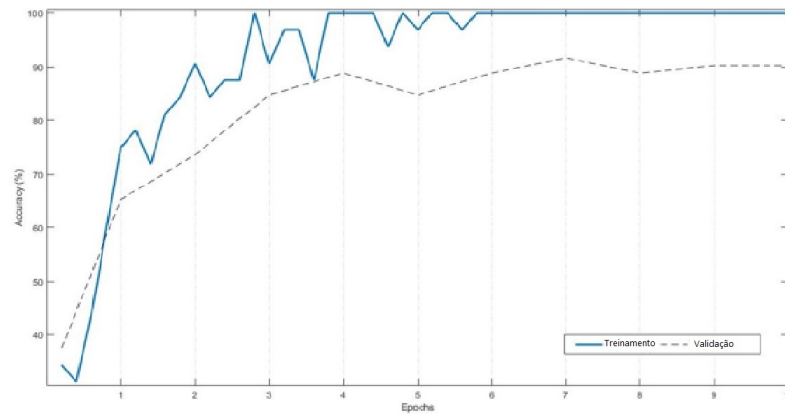


Figura 27 – Desempenho do *fold* 1 com representação GR durante o treinamento e validação da rede EfficientNet *baseline*.

Fonte: Elaborado pelo autor.

Nas Figuras 29 e 28, são apresentadas as matrizes de confusão do *fold* 1 para as representações com a EfficientNet e a ResNet-50. Nestas figuras, a classe 1 representa a imagem saudável, a classe 2 define a displasia leve, a classe 3 representa a displasia moderada e a classe 4 representa a displasia severa. Em todos os casos, é possível observar uma alta taxa de acerto para a classe 4 (severa), independente da representação ou modelo de rede empregada. Com a relação a classificação com as imagens originais, A) nota-se uma taxa de erro maior entre as classes 1 e 2 (saudável e leve) e o mesmo ocorre no caso da representação GR. Ao observar as demais representações, tem-se uma distribuição bem maior, mostrando que houve um número maior de erro de classificação, o que justifica os valores de acurácia encontrados nas Tabelas 2. Essas matrizes ainda mostram que em ambas as redes a classe severa obteve somente valores verdadeiros positivos para a representação GR.

Uma avaliação estatística foi realizada com a ANOVA de medições repetidas comparando as representações para avaliar a relevância dos resultados. Neste teste, cada *fold* foi considerado como uma medição repetida. Os resultados mostraram um p -valor igual ou menor que 0,001. Com isso, se nota uma diferença estatística significativa entre as representações investigadas. Com o objetivo de mostrar em quais grupos essas diferenças ocorreram, então, foi realizado o teste *post-hoc* Tuckey comparando cada um dos modelos. A Tabela 3 apresenta os resultados da comparação entre os modelos. Para todas as análises (pares de representações), o valor de p -valor é inferior a 0,05, o que permite afirmar que os resultados obtidos para todos os modelos são distintos estatisticamente. Numa avaliação dos resultados com a métrica acurácia, é possível notar que a representação GR traz os melhores resultados.

Na Figura 30, é apresentado estimativa das marginais do teste realizado com intervalo de confiança de 95%. No gráfico, o eixo y representa a distribuição da variável dependente (acurácia), enquanto o eixo x apresenta o modelo de representação. Esse figura traz

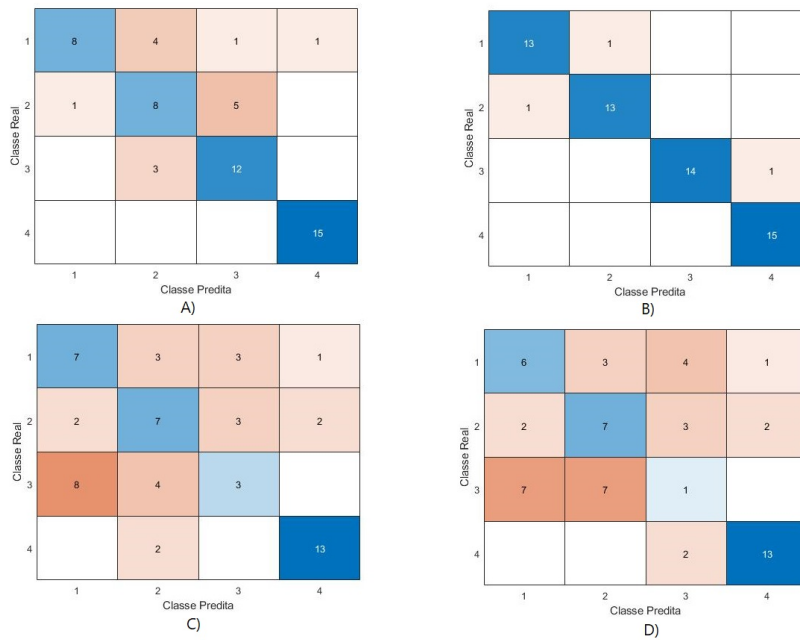


Figura 28 – Matriz de confusão para EfficientNet com as representações investigadas: A) O, B) GR, C) MTF e D) SSM,
Fonte: Elaborado pelo autor.

Tabela 3 – Resultados dos ρ -valores do teste *post – hoc* Tuckey para as representações investigadas independente da estrutura da rede.

	O	GR	SSM	MTF
O	-	0,001	<0,001	<0,001
GR	0,001	-	<0,001	<0,001
SSM	<0,001	<0,001	-	0,009
MTF	<0,001	<0,001	0,009	-

visualmente a diferença estatística entre as representações encontradas no teste e destaca a relevância da representação GR em relação as outras.

Um segundo teste foi realizado com o objetivo de identificar se há uma diferença significativa ao comparar o resultado entre as estruturas das CNNs. A medida de ρ -valor foi de 0,976, o que permite afirmar que não há uma diferença estatística relevante entre as estruturas das CNNs. Esse comportamento pode ser observado na Figura 31.

Com objetivo de uma avaliação mais detalhada das classes, as Tabelas 4 e 5 trazem o comportamento dos modelos para cada uma das classes de imagens analisadas. De forma semelhante aos dados apresentados na Tabela 2, a representação GR proporcionou os melhores valores da métrica acurácia para cada classe investigada. Analisando a Tabela 4, nota-se que para algumas classes, saudável e displasia leve, os valores de acurácia foram inferiores a 96% no processo de classificação com GR. Quando se avalia as lesões nos estágios mais avançados, displasia moderada e displasia severa, os valores da métrica acurácia foram mais relevantes na representação GR. Isso mostra que a estratégia é mais

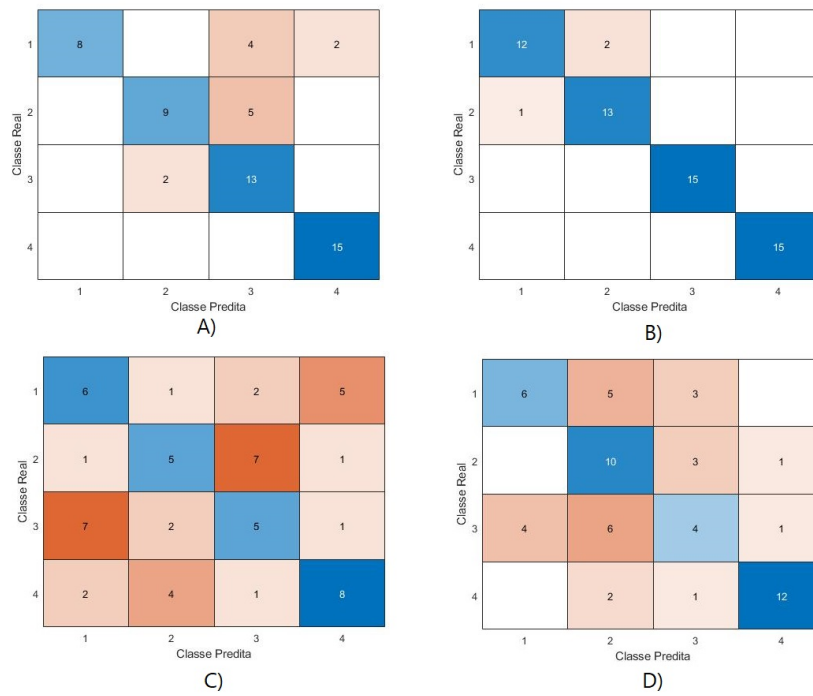


Figura 29 – Matriz de confusão da rede ResNet-50 para cada uma das representações investigadas: A) O, B) GR, C) MTF e D) SSM.

Fonte: Elaborado pelo autor.

efetiva para os estágios mais avançados da doença. Em relação as demais métricas, nesse modelo, apenas as displasias leves tiveram uma sensibilidade inferior a 90% usando a representação GR. O modelo GR permite elevar os valores de sensibilidade para as diversas classes de imagens histológicas. Com o modelo da ResNet, tem-se um melhoramento em relação aos valores da imagem original (ver Tabela 5) com valores de acurácia acima de 93% para a displasia leve. No entanto, com esse modelo de rede, há uma queda nos valores de sensibilidade para algumas classes de lesões.

5.2 Arquiteturas com Otimização das Camadas Totalmente Conectadas

A Tabela 6 apresenta os resultados de classificação com a associação das CNNs e as representações, mas com as camadas totalmente conectadas otimizadas. Os melhores resultados médios obtidos foram destacados em negrito para cada uma das arquiteturas. Ao compará-las com a Tabela 2, é possível observar que independente da representação, houve um aumento nos valores dos resultados da acurácia para a arquitetura EfficientNet, na qual para imagem original o valor de acurácia aumentou de 90,2% para 93,4%. Com uso da representação SSM, os valores de acurácia foram modificados de 75,8% para 78,0%, com o modelo MTF, os valores foram de 72,8% para 76,8%. Finalmente, para a representação GR esse valor foi de 96,6% para 98,6%, proporcionando um aumento médio de 2,8%.

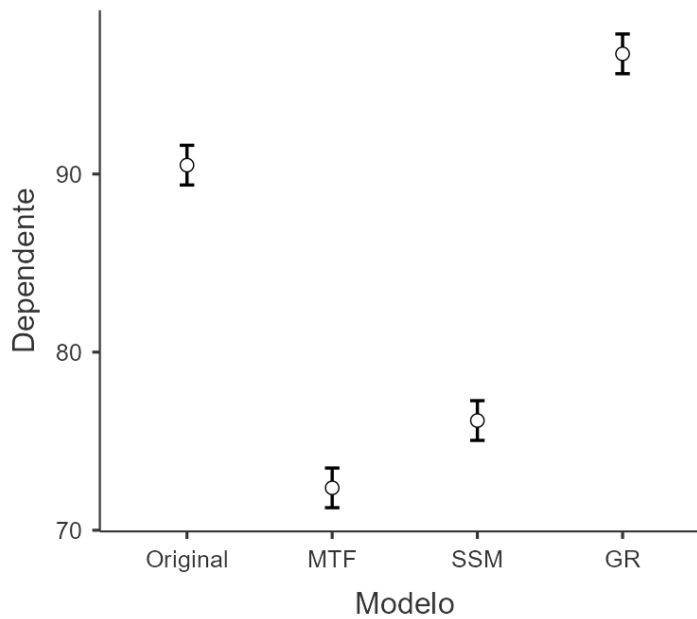


Figura 30 – Médias marginais estimadas para os modelos *baseline*, em que gráfico de dispersão é a estimativa do teste estatístico.

Fonte: Elaborado pelo autor.

Tabela 4 – Valores médios obtidos com as métricas Acurácia (Acc), Sensibilidade (Sens) e Especificidade (Esp) para cada uma das classes de imagens investigadas da cavidade oral com a arquitetura EfficientNet.

		Saudável	Leve	Moderada	Severa
O	Acc	89,8 ± 02,2	85,1 ± 04,6	87,1 ± 02,7	98,6 ± 01,4
	Sens	79,4 ± 15,2	71,4 ± 12,7	71,3 ± 13,2	98,7 ± 03,0
	Esp	93,2 ± 04,8	89,6 ± 04,8	92,3 ± 05,7	98,6 ± 01,2
MTF	Acc	74,2 ± 06,0	72,3 ± 03,8	72,6 ± 05,3	78,1 ± 05,0
	Sens	47,3 ± 18,3	37,9 ± 10,4	40,8 ± 11,8	68,8 ± 08,5
	Esp	83,3 ± 04,8	83,8 ± 03,1	83,3 ± 09,8	81,2 ± 08,7
SSM	Acc	78,6 ± 07,7	71,3 ± 04,5	69,5 ± 08,4	83,8 ± 04,7
	Sens	53,9 ± 15,1	38,0 ± 18,5	35,1 ± 19,7	79,7 ± 10,6
	Esp	86,9 ± 07,2	82,4 ± 04,0	81,0 ± 08,4	85,2 ± 05,0
GR	Acc	95,2 ± 03,3	94,9 ± 04,1	98,3 ± 01,2	98,0 ± 01,9
	Sens	95,9 ± 03,7	85,2 ± 15,9	96,0 ± 03,7	95,8 ± 06,3
	Esp	95,0 ± 04,2	98,2 ± 01,9	99,1 ± 02,0	98,6 ± 01,2

Esse comportamento também foi observado na arquitetura ResNet-50, mas o aumento foi menor em relação aos valores obtidos com a rede EfficientNet. Para esse modelo, o valor de acurácia para as imagens originais foram de 90,9% para 91,1%, a representação SSM foi de 76,7% para 77,0%, o modelo MTF foi de 71,9% para 77,6% e com a representação GR esse aumento elevou o valor de 96,9% para 98,0%, um aumento médio nas representações de 1,3%.

Nas Figuras 33 e 32, são apresentadas as matrizes de confusão com dados do *fold*

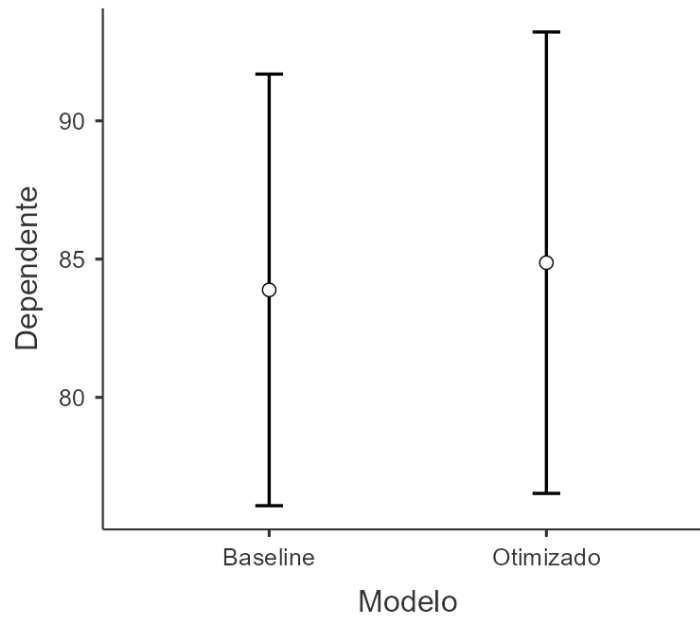


Figura 31 – Valores de médias marginais estimadas para as estruturas das CNNs em que gráfico de dispersão é a estimativa do teste estatístico.

Fonte: Elaborado pelo autor.

Tabela 5 – Valores médios obtidos com as métricas acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para cada uma das classes de imagens investigadas da cavidade oral com a arquitetura ResNet-50.

		Saudável	Leve	Moderada	Severa
O	Acc	91,9 ± 01,5	87,8 ± 03,7	85,5 ± 02,9	98,3 ± 01,7
	Sens	82,1 ± 14,2	72,9 ± 11,0	72,9 ± 08,7	98,7 ± 03,0
	Esp	95,0 ± 02,9	92,8 ± 03,1	89,6 ± 06,4	98,2 ± 01,9
MTF	Acc	72,6 ± 04,2	69,9 ± 05,4	69,9 ± 05,7	75,4 ± 07,3
	Sens	43,2 ± 03,3	32,5 ± 13,8	40,5 ± 13,1	59,4 ± 06,8
	Esp	82,4 ± 05,7	82,4 ± 03,3	79,7 ± 09,3	80,7 ± 11,4
SSM	Acc	80,4 ± 04,2	71,3 ± 03,3	73,6 ± 03,3	81,5 ± 08,6
	Sens	56,6 ± 18,6	43,6 ± 17,2	37,8 ± 07,4	75,6 ± 10,3
	Esp	88,2 ± 07,1	80,6 ± 06,6	85,6 ± 03,4	83,5 ± 11,0
GR	Acc	94,2 ± 03,6	93,9 ± 03,4	99,7 ± 00,7	100,0 ± 0,0
	Sens	87,8 ± 03,1	89,2 ± 13,0	98,7 ± 03,0	100,0 ± 00,0
	Esp	96,3 ± 04,6	95,5 ± 01,6	100,0 ± 00,0	100,0 ± 00,0

1 para as imagens O A), representação GR B), representação MTF C) e representação SSM D) com as arquiteturas EfficientNet e ResNet-50, respectivamente. O *fold* 1 foi escolhido em ambos os modelos (*baseline* e otimizado). Semelhante ao modelos *baselines*, a classe 1 representa a imagem saudável, a classe 2 representa a lesão leve, a classe 3 representa a displasia moderada e a classe 4 representa a displasia severa. Nestas matrizes, independente da rede, se nota o melhor desempenho com a classificação das lesões da classe displasia severa para os modelos de representação O e GR. Nesse experimento,

Tabela 6 – Valores de acurácia para as arquiteturas com otimização das camadas totalmente conectadas.

EfficientNet						
	<i>Fold 1</i>	<i>Fold 2</i>	<i>Fold 3</i>	<i>Fold 4</i>	<i>Fold 5</i>	Média (σ)
O	92,2	93,1	95,8	93,3	92,5	93,4 $\pm$ 1,4
MTF	75,9	74,1	76,7	76,7	80,8	76,8 $\pm$ 2,5
SSM	75,9	78,4	77,5	79,2	79,2	78,0 $\pm$ 1,4
GR	100,0	96,6	99,2	98,3	99,2	98,6 $\pm$ 1,3

ResNet-50						
	<i>Fold 1</i>	<i>Fold 2</i>	<i>Fold 3</i>	<i>Fold 4</i>	<i>Fold 5</i>	Média (σ)
O	92,2	92,2	92,5	86,7	91,7	91,1 $\pm$ 2,5
MTF	77,6	74,1	75,8	78,3	79,2	77,0 $\pm$ 2,0
SSM	79,3	78,4	75,8	79,2	75,0	77,6 $\pm$ 2,0
GR	100,0	95,7	96,7	99,2	98,3	98,0 $\pm$ 1,8

ainda é possível notar que a classificação com o GR permitiu obter resultados relevantes com todas as classes após a otimização das camadas totalmente conectadas. As demais representações apresentaram melhoramento para alguns grupos de lesões.

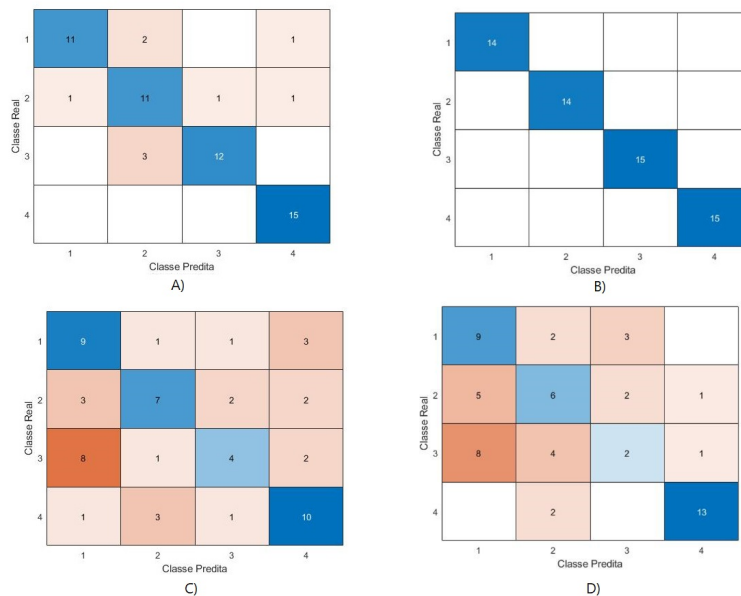


Figura 32 – Matrizes de confusão da rede otimizada EfficientNet para as representações investigadas: (A) O, (B) GR, (C) MTF e (D) SSM.

Fonte: Elaborado pelo autor.

As Tabelas 7 e 8 apresentam os resultados de acurácia (Acc) média dos *folds*, bem como os valores de sensibilidade (Sens) e especificidade (Esp) para a estruturas investigadas para cada uma das classes. Ao analisar os resultados em relação as Tabelas 4 e 5, observa-se que na maioria dos casos, independente da rede, houve um incremento no valor médio da acurácia. Um exemplo desse comportamento é observado com a imagem O em que o valor sem uso da otimização era de 91,9% e foi aprimorado para 93,2% para a classe saudável

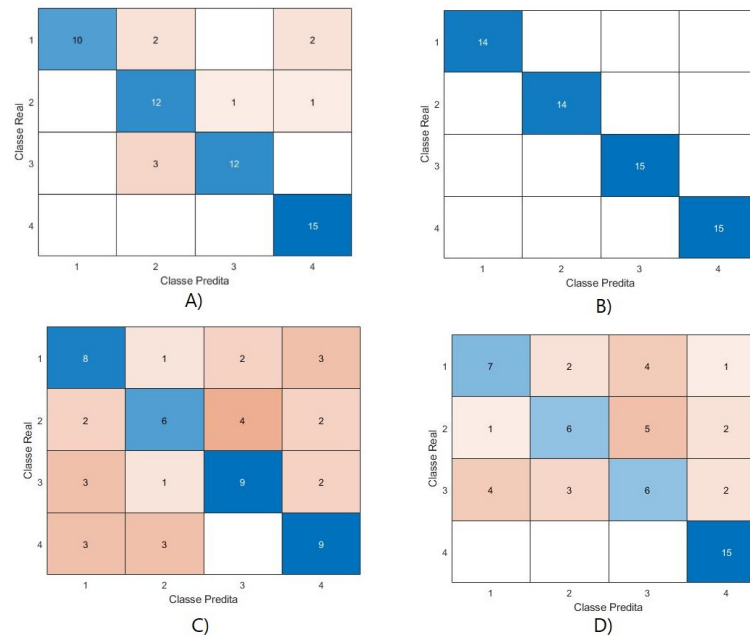


Figura 33 – Matrizes de confusão da rede otimizada ResNet-50 para as representações investigadas: (A) O, (B) GR, (C) MTF e (D) SSM.

Fonte: Elaborado pelo autor.

com a estrutura da ResNet-50. Semelhante as avaliações anteriores, a representação GR manteve o melhor desempenho para os grupos investigados. O comportamento dos dados com as representações MTF e SSM continuaram inferiores aos valores obtidos com as imagens O e GR.

Tabela 7 – Resultados experimentais de acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para as classes investigadas com otimização das camadas totalmente conectadas com estrutura baseada em ResNet-50.

		Saudável	Leve	Moderada	Severa
O	Acc	93,2 ± 1,20	85,5 ± 5,00	88,60 ± 2,90	97,0 ± 3,60
	Sens	82,3 ± 8,20	75,8 ± 9,90	73,0 ± 12,3	97,3 ± 6,00
	Esp	96,8 ± 2,00	88,8 ± 5,40	93,7 ± 4,30	96,8 ± 3,50
MTF	Acc	80,3 ± 5,20	72,3 ± 3,70	75,3 ± 4,10	80,1 ± 3,60
	Sens	60,8 ± 3,50	47,2 ± 10,2	43,0 ± 16,3	65,0 ± 6,80
	Esp	86,8 ± 6,70	80,7 ± 4,60	86,0 ± 1,90	85,1 ± 2,50
SSM	Acc	80,7 ± 4,50	74,4 ± 5,10	74,0 ± 4,20	81,1 ± 6,70
	Sens	59,3 ± 15,3	28,6 ± 8,50	49,9 ± 8,30	82,4 ± 12,2
	Esp	87,8 ± 6,00	89,7 ± 7,10	82,0 ± 4,20	80,7 ± 4,90
GR	Acc	98,0 ± 1,80	95,9 ± 3,50	100,0 ± 0,00	100,0 ± 0,00
	Sens	89,3 ± 10,1	94,7 ± 5,60	100,0 ± 0,00	100,0 ± 0,00
	Esp	98,2 ± 1,90	96,4 ± 3,40	100,0 ± 0,00	100,0 ± 0,00

Para analisar a relevância estatística, o teste ANOVA de medições repetidas foi em-

Tabela 8 – Resultados experimentais com valores da acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para as classes investigadas com otimização das camadas totalmente conectadas com estrutura baseada em EfficientNet.

		Saudável	Leve	Moderada	Severa
O	Acc	94,6 ± 2,60	89,5 ± 2,60	92,2 ± 2,50	97,3 ± 0,90
	Sens	85,0 ± 10,1	83,7 ± 4,10	79,8 ± 4,40	98,7 ± 1,70
	Esp	97,7 ± 0,00	91,4 ± 2,50	96,4 ± 2,50	96,8 ± 1,30
MTF	Acc	77,6 ± 5,80	74,0 ± 3,50	75,9 ± 6,50	79,7 ± 1,80
	Sens	56,9 ± 10,7	44,7 ± 10,4	40,6 ± 14,2	73,0 ± 13,2
	Esp	84,7 ± 7,60	83,8 ± 3,80	87,8 ± 8,80	82,0 ± 5,10
SSM	Acc	77,0 ± 6,40	73,6 ± 2,60	74,3 ± 3,40	87,2 ± 4,70
	Sens	58,2 ± 18,3	37,9 ± 8,00	44,3 ± 14,7	85,0 ± 5,90
	Esp	83,3 ± 9,20	85,6 ± 3,00	84,7 ± 3,60	87,9 ± 7,40
GR	Acc	98,3 ± 2,10	98,0 ± 1,90	99,0 ± 0,90	99,3 ± 0,90
	Sens	96,0 ± 8,90	97,3 ± 3,70	97,3 ± 3,70	98,6 ± 3,20
	Esp	99,1 ± 1,20	98,2 ± 3,00	99,5 ± 1,00	99,6 ± 1,00

pregado, utilizando as redes otimizadas e não otimizadas como modelos distintos. O teste ANOVA trouxe um ρ -valor $< 0,001$, o que indica que há diferença significativa entre as representações investigadas. Assim, para identificar em quais grupos ocorre a diferença, foi realizado um teste *post hoc* comparando cada modelo entre si. A Tabela 9 apresenta esse resultado em que o ρ -valor $< 0,05$ indica uma diferença significativa entre os modelos. Através destes resultados, destacados em negrito, pode-se identificar que para os modelos, O e MTF, a otimização da estrutura das redes provocou um aumento estatisticamente relevante, enquanto que para o SSM e o GR, apesar do aumento absoluto de acurácia visto nos resultados anteriores, esse aumento não foi estatisticamente distinto.

Tabela 9 – Resultados de ρ -valor do teste *post hoc* para as representações investigadas comparando os valores não otimizados e otimizados, independente da estrutura da rede.

	Otimizado			
	O	GR	SSM	MTF
O	0,05	$<0,001$	$<0,001$	$<0,001$
GR	0,003	0,407	$<0,001$	$<0,001$
SSM	$<0,001$	$<0,001$	0,092	0,065
MTF	$<0,001$	$<0,001$	$<0,001$	0,001

As Tabelas 11 e 10 mostram os parâmetros dos modelos em cada *fold* durante a otimização das camadas totalmente conectadas dos modelos das arquiteturas EfficientNet e ResNet-50, respectivamente, conforme descrito na Seção 4.6. Esta etapa demandou alto tempo computacional, uma vez que para cada conjunto (*fold*), o método exigiu em torno de 120 minutos para obter as 10 gerações. O critério de parada foi adotado 10 gerações baseado no estudo de (GONÇALVES; SOUZA; FERNANDES, 2022). Nas Tabelas 11

e 10, o parâmetro $N1$ representa os números de neurônios da primeira camada totalmente conectada, o $N2$ representa os números de neurônios da segunda camada totalmente conectada, no qual a coluna $N2$ só apresenta valor quando a camada é considerada no modelo. Nestes modelos, ainda há uma camada onde está presente a camada softmax e não foi considerada no processo de otimização. Já o gene Dr representa a taxa de *dropout*, enquanto os parâmetros *peso* e *bias* representam, respectivamente, os tipos de inicialização do peso e bias. Por fim, o termo Ta representa a taxa aprendizado inicial do modelo durante o treinamento.

Em relação ao tamanho das estruturas totalmente conectadas é possível observar que para a ResNet-50 (ver Tabela 11), a estrutura de 3 camadas foi considerada na maior parte dos conjuntos durante a avaliação do modelo de representação SSM. Esse comportamento também foi observado com a EfficientNet (ver Tabela 10). Apesar de um número maior de camadas, o modelo não trouxe resultados relevantes. Em relação ao gene Dr , independente da estrutura e representação analisada, é possível verificar que, na grande maioria dos *folds*, é necessário uma taxa maior que 20% para otimizar a estrutura. Essa camada ajuda a melhorar a capacidade de generalização da rede, tornando o modelo mais robusto e eficiente, uma vez que uma fração aleatória dos neurônios é temporariamente desativada durante o treinamento, o que ajuda a reduzir o *overfitting*. Além disso, é observado que o tipo mais recorrente de inicialização bias para o caso estudado foi 1 - zeros, em que a matriz de bias é inicializada com valores configurados em zero. Essa matriz é um componente fundamental em redes neurais, pois permite que a rede neural aprenda a ajustar os valores de saída independentemente das entradas. Esse ajuste é útil para introduzir deslocamentos nos padrões de entrada e para permitir que a rede capture relações mais complexas nos dados.

Para a inicialização dos pesos, o valor 3, que representa *orthogonal*, foi o que teve a maior ocorrência para a ResNet-50 (12/20), enquanto que para a rede EfficientNet, essa inicialização ocorreu mais com o valor 1, *glorot*, e valor 2, *he*. O modelo 1, *glorot*, visa manter a variância constante durante a propagação, mantendo assim um fluxo de gradiente estável durante o treinamento, o que ajuda a melhorar a convergência do treinamento e o desempenho geral do modelo. O modelo 2, *he*, permite uma propagação mais eficaz de gradientes durante o treinamento ao ser focada para lidar com funções de ativação retificadoras da ReLU, que estão presentes tanto na arquitetura da ResNet-50 quanto da EfficientNet. Por fim, o modelo 3, *orthogonal*, ajuda a garantir que os gradientes durante o treinamento não sejam amplificados ou atenuados de forma significativa, evitando problemas de desaparecimento de gradientes.

Em relação a taxa de aprendizado inicial, gene Ta , é possível notar que com a ResNet-50, seu valor foi bem menor do que em relação ao EfficientNet, uma vez que esse hiperparâmetro controla o tamanho das atualizações dos pesos durante o treinamento de um modelo. Assim, pelos dados de Ta , a EfficientNet apresenta uma valor inicial maior, que

pode resultar em grandes oscilações nos valores dos pesos, e dificulta a convergência do modelo para uma solução ótima, contudo, como há um decaimento da taxa de aprendizado ao longo do treinamento, esse problema pode ser contornado. Já a ResNet-50 apresentou valores mais baixos, o que representa uma convergência suave, evitando oscilações excessivas com um treinamento mais estável.

Tabela 10 – Avaliação do modelo EfficientNet com a otimização das camadas totalmente conectadas com uso do AG.

Representação		N1	N2	Dr	Bias	Peso	Ta
O	<i>Fold1</i>	2^8	2^{10}	20,0	2	1	10^{-2}
	<i>Fold2</i>	2^5	-	6,7	1	3	10^{-4}
	<i>Fold3</i>	2^6	2^4	33,3	1	1	10^{-1}
	<i>Fold4</i>	2^9	-	46,7	2	2	10^{-1}
	<i>Fold5</i>	2^9	2^4	40,0	1	1	10^{-2}
MTF	<i>Fold1</i>	2^{10}	2^9	20,0	1	1	10^{-1}
	<i>Fold2</i>	2^9	2^3	53,3	1	2	10^{-3}
	<i>Fold3</i>	2^5	-	33,3	3	3	10^{-1}
	<i>Fold4</i>	2^7	-	6,7	1	2	10^{-1}
	<i>Fold5</i>	2^8	2^8	40,0	1	3	10^{-5}
SSM	<i>Fold1</i>	2^3	2^6	33,3	3	2	10^{-6}
	<i>Fold2</i>	2^4	2^4	0,0	3	1	10^{-2}
	<i>Fold3</i>	2^3	2^7	6,7	2	1	10^{-4}
	<i>Fold4</i>	2^7	2^4	51,67	3	2	10^{-5}
	<i>Fold5</i>	2^{10}	2^6	13,3	1	3	10^{-4}
GR	<i>Fold1</i>	2^{10}	-	60,0	1	2	10^{-1}
	<i>Fold2</i>	2^{10}	-	60,0	2	1	10^{-5}
	<i>Fold3</i>	2^9	2^3	33,3	1	2	10^{-6}
	<i>Fold4</i>	2^5	2^3	6,7	1	2	10^{-2}
	<i>Fold5</i>	2^{10}	2^9	60,0	1	1	10^{-1}

Na Figura 34, é apresentado o histograma da distribuição da quantidade de neurônios na primeira camada totalmente conectada (gene *N1*). As barras na cor laranja representam a quantidade de vezes que aquele valor foi obtido com a estrutura da EfficientNet. As barras na cor azul representam a quantidade de vezes para estrutura com a ResNet-50. É possível observar que para cada estrutura de CNN, a distribuição da quantidade de neurônios na primeira camada totalmente conectada foi diferente. Para ResNet-50, a maioria dos resultados (13/20) se concentraram entre 2^5 e 2^7 neurônios, enquanto que para a EfficientNet, as maiores ocorrências foram com 2^9 e 2^{10} neurônios. Isso demonstra que, para ambas as estruturas, um maior número de neurônios é promissor. Isso ocorre porque o número de neurônios em cada camada totalmente conectada está diretamente ligado à capacidade do modelo de aprender representações complexas dos dados, auxiliando na captura de padrões mais sutis nos dados.

Tabela 11 – Resultados da otimização das camadas totalmente conectadas com AG para o modelo ResNet-50.

Representação		N1	N2	Dr (%)	Bias	Peso	Ta
O	<i>Fold1</i>	2 ⁵	2 ⁶	6,7	1	1	10 ⁻³
	<i>Fold2</i>	2 ³	-	26,7	1	3	10 ⁻¹
	<i>Fold3</i>	2 ⁵	-	53,3	3	4	10 ⁻⁵
	<i>Fold4</i>	2 ⁶	-	40,0	3	1	10 ⁻⁵
	<i>Fold5</i>	2 ⁶	-	40,0	2	3	10 ⁻⁴
MTF	<i>Fold1</i>	2 ⁷	-	6,7	2	2	10 ⁻³
	<i>Fold2</i>	2 ⁵	-	33,3	1	2	10 ⁻⁵
	<i>Fold3</i>	2 ⁸	-	53,3	1	2	10 ⁻³
	<i>Fold4</i>	2 ⁹	2 ⁶	20,0	1	3	10 ⁻³
	<i>Fold5</i>	2 ⁷	2 ⁴	20,0	3	3	10 ⁻¹
SSM	<i>Fold1</i>	2 ⁶	2 ⁹	60,0	3	3	10 ⁻²
	<i>Fold2</i>	2 ³	2 ⁹	26,7	3	1	10 ⁻³
	<i>Fold3</i>	2 ⁸	-	26,7	1	4	10 ⁻⁵
	<i>Fold4</i>	2 ⁷	2 ⁸	13,3	3	3	10 ⁻²
	<i>Fold5</i>	2 ⁸	2 ¹⁰	0,0	1	3	10 ⁻³
GR	<i>Fold1</i>	2 ⁹	-	46,7	1	3	10 ⁻⁴
	<i>Fold2</i>	2 ⁶	-	13,3	3	3	10 ⁻²
	<i>Fold3</i>	2 ⁶	2 ⁴	0,0	1	3	10 ⁻⁵
	<i>Fold4</i>	2 ⁷	2 ⁶	66,7	1	3	10 ⁻¹
	<i>Fold5</i>	2 ⁵	-	13,3	2	3	10 ⁻³

5.3 Comitê de Classificadores

Após realizado os experimentos com as arquiteturas CNNs e as representações, essa etapa traz os comitês de composição das CNNS com regra da soma e pesos iguais. Essa abordagem de parâmetros segue o modelo adotado por Roberto, Neves, Lumini, Martins e Nascimento (2024). Conforme descrito na Seção 4.7, os comitês utilizados foram os homogêneos, em que todos os classificadores apresentam a mesma estrutura. As variações ocorreram com as imagens de representação para as redes, sendo o modelo 1 - imagens originais e representação de GR (O + GR), modelo 2 - imagens originais e representação MTF (O + MTF), modelo 3 - imagens originais e representação SSM (O + SSM). Com base nos resultados anteriores, a representou GR foi investigada no modelo 4 - representações GR e SSM (GR + SSM), modelo 5 - representações GR e MTF (GR + MTF) e o modelo 6 - representações (GR + MTF + SSM). Finalmente, uma composição de todas as entradas, modelo 7 - (O + GR + MTF + SSM), foi investigada.

5.3.1 Comitê com as Estruturas das CNNs

A Tabela 12 apresenta os resultados da classificação para cada *fold* com as redes investigadas e composição das representações. Ao analisar os resultados se observa que os comitês com as imagens originais e a representação GR apresentaram melhores valores

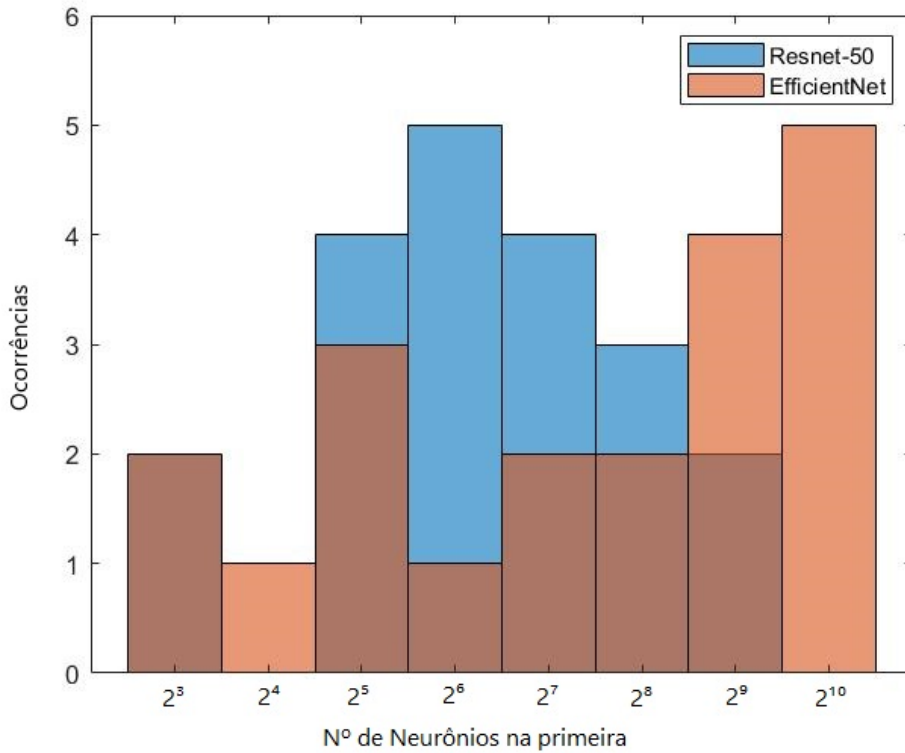


Figura 34 – Histograma da distribuição do número de neurônios na primeira camada totalmente conectada, em cor laranja, para a EfficientNet, e em cor azul para a Resnet-50. **Fonte:** Elaborado pelo autor.

para ambas arquiteturas CNNs (acurácia de 98,1%). Quando ocorreu a associação da imagem original com as representações MTF e SSM (modelo 2 e modelo 3), é possível notar que os valores de acurácia pioram em relação ao modelo 1. Com as associações entre a representação GR e com as demais (modelos 4 e 5), para ambas as redes, os valores obtidos foram entre 93% e 95%. As abordagens que exploram composições superiores a duas técnicas (modelos 6 e 7) não proporcionaram resultados melhores.

Com objetivo de verificar a relevância estatística dos resultados, o teste ANOVA de medições repetidas foi empregado aos *folds*. Os valores com relevância estatística estão marcados em negrito. O teste resultou um valor de p -valor de 0,002, o que indica que há diferença estatística entre os comitês. Assim, para identificar em quais grupos estão os valores, foi realizado um teste *post hoc* Tukey comparando os modelos. A Tabela 13 apresenta os resultados para cada um dos modelos propostos. Por meio desses resultados, se pode identificar que há, para algumas associações, uma diferença significativa, por exemplo, o comitê modelo 1 em relação aos modelos 2 e 3. Contudo, ao analisar esse comitê as demais associações do modelo GR e as outras representações, o valor de p -valor é superior a 0,05, ou seja, não há relevância estatística. Quando ocorreu a composição de três representações, o teste mostrou que há relevância estatística entre o modelo 1 e o modelo 6 (p -valor = 0,004), ou seja, o modelo 1 é melhor do que a composição das três representações. Na Figura 35, é apresentado os intervalos dessa análise em que o modelo

Tabela 12 – Desempenho dos modelos propostos com comitês das CNNs e representações extraídas de descritores fractais sem o processo de otimização das camadas totalmente conectadas

EfficientNet						
	<i>Fold 1</i>	<i>Fold 2</i>	<i>Fold 3</i>	<i>Fold 4</i>	<i>Fold 5</i>	Média (σ)
modelo 1	99,1	96,5	99,2	97,5	98,3	98,1 $\pm$ 1,1
modelo 2	87,9	86,2	89,2	91,2	91,2	89,3 $\pm$ 2,4
modelo 3	86,2	83,6	88,3	89,1	91,6	87,8 $\pm$ 3,0
modelo 4	95,6	91,2	95,0	91,6	96,6	94,1 $\pm$ 2,4
modelo 5	97,4	90,5	97,5	95,8	97,5	95,8 $\pm$ 3,0
modelo 6	90,5	81,8	90,1	90,0	94,1	89,5 $\pm$ 4,5
modelo 7	97,4	88,7	95,0	95,0	97,5	94,9 $\pm$ 2,1

ResNet-50						
	<i>Fold 1</i>	<i>Fold 2</i>	<i>Fold 3</i>	<i>Fold 4</i>	<i>Fold 5</i>	Média (σ)
modelo 1	97,4	96,5	99,2	99,2	99,2	98,1 $\pm$ 1,1
modelo 2	88,7	83,6	91,2	84,2	88,3	87,3 $\pm$ 3,4
modelo 3	89,6	88,8	91,7	89,2	90,0	89,9 $\pm$ 1,1
modelo 4	97,5	93,1	97,5	91,8	97,5	95,4 $\pm$ 2,8
modelo 5	97,4	90,5	95,0	89,2	96,7	93,8 $\pm$ 3,7
modelo 6	93,4	84,5	90,8	81,7	91,7	88,5 $\pm$ 5,2
modelo 7	94,4	88,8	95,0	90,8	95,0	92,9 $\pm$ 2,9

1, representado pelo rótulo 1, é o valor mais relevante dos comitês estudados. Contudo, é possível observar que há uma sobreposição da estimativa entre os modelos 1, 4 e 5, evidenciando a sobreposição do desvios, os quais indicam a não relevância estatística.

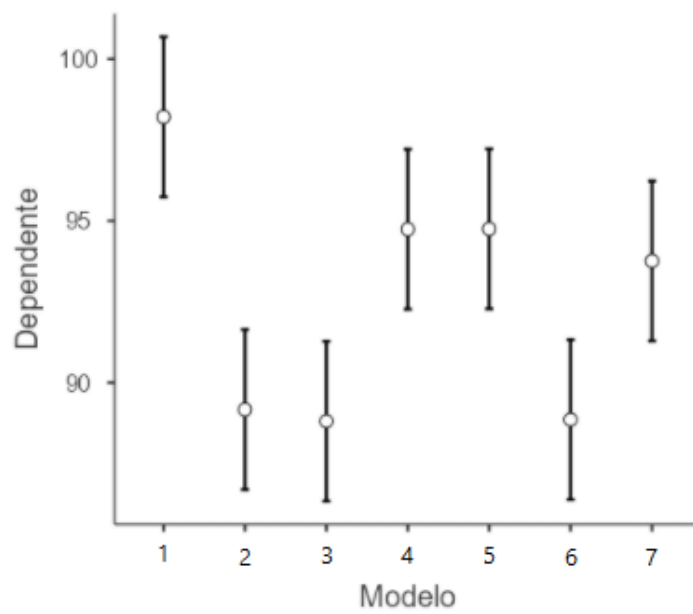


Figura 35 – Médias marginais estimadas para as estruturas dos comitês investigados, em que gráfico de dispersão é a estimativa do teste estatístico. **Fonte:** Elaborado pelo autor.

Tabela 13 – Resultados dos ρ -valores do teste *post-hoc* para os comitês investigados independente da estrutura da rede.

	modelo 1	modelo 2	modelo 3	modelo 4	modelo 5	modelo 6	modelo 7
modelo 1	-	0,005	0,004	0,336	0,339	0,004	0,155
modelo 2	0,005	-	1,000	0,063	0,062	1,000	0,139
modelo 3	0,004	1,000	-	0,047	0,047	1,000	0,103
modelo 4	0,336	0,063	0,047	-	1,000	0,049	0,991
modelo 5	0,339	0,062	0,047	1,000	-	0,049	0,991
modelo 6	0,004	1,000	1,000	0,049	0,049	-	0,108
modelo 7	0,155	0,139	0,103	0,991	0,991	0,108	-

As Tabelas 14 e 15 apresentam os valores médios de acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para cada classe, obtidos com as abordagens de comitês para as redes EfficientNet e ResNet-50, respectivamente. O mesmo padrão pode ser identificado na qual, independente da associação das representações e dos modelos de redes, os valores de acurácia para a classe severa são os mais elevados. Em algumas classes nota-se um aumento em relação ao *baseline* (ver Tabelas 4 e 5). Um exemplo desse comportamento é os valores da métrica acurácia para a classe leve usando a ResNet-50, em que na avaliação apenas a representação GR, que tinha valor de Acc de 93,9%, e os resultados com o comitê modelo 1 proporcionou um valor de 96,6%. Os demais valores das medidas foram aprimorados, como no caso da medida Sens, para as amostras saudáveis com a representação GR, em que no baseline era de 82,1% e com o uso do comitê modelo 1 esse valor foi para 94,5%.

Tabela 14 – Resultados experimentais de acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para as classes investigadas com o modelos de comitês e estrutura baseada em ResNet-50.

Comitê	Métrica	Saudável	Leve	Moderado	Severa
modelo 1	Acc	$97,3 \pm 2,6$	$96,6 \pm 3,4$	$99,0 \pm 0,9$	$99,7 \pm 0,8$
	Sens	$94,5 \pm 5,9$	$91,9 \pm 7,3$	$98,7 \pm 3,0$	$100,0 \pm 0,0$
	Esp	$98,2 \pm 3,0$	$98,2 \pm 1,9$	$99,1 \pm 1,2$	$99,5 \pm 1,0$
modelo 2	Acc	$89,8 \pm 2,5$	$82,7 \pm 7,6$	$82,1 \pm 5,6$	$94,6 \pm 2,2$
	Sens	$77,0 \pm 10,3$	$59,4 \pm 17,0$	$66,2 \pm 12,5$	$96,0 \pm 3,7$
	Esp	$94,1 \pm 2,75$	$90,5 \pm 5,8$	$87,3 \pm 6,9$	$94,2 \pm 2,5$
modelo 3	Acc	$92,2 \pm 2,0$	$85,4 \pm 4,0$	$85,1 \pm 1,9$	$96,6 \pm 2,6$
	Sens	$82,2 \pm 10,0$	$66,2 \pm 14,2$	$71,7 \pm 10,7$	$98,7 \pm 3,0$
	Esp	$95,5 \pm 3,6$	$91,9 \pm 1,9$	$89,6 \pm 4,6$	$96,0 \pm 2,9$
modelo 4	Acc	$93,6 \pm 3,4$	$92,6 \pm 4,5$	$98,3 \pm 1,7$	$97,3 \pm 5,1$
	Sens	$85,1 \pm 2,9$	$82,6 \pm 17,9$	$96,0 \pm 6,0$	$100,0 \pm 0,0$
	Esp	$96,4 \pm 3,5$	$95,9 \pm 1,9$	$99,1 \pm 2,0$	$96,4 \pm 6,8$
modelo 5	Acc	$93,9 \pm 3,6$	$90,2 \pm 6,8$	$94,6 \pm 4,3$	$96,3 \pm 2,7$
	Sens	$83,5 \pm 3,5$	$77,2 \pm 22,3$	$91,9 \pm 5,6$	$97,3 \pm 3,7$
	Esp	$97,3 \pm 3,8$	$94,6 \pm 2,0$	$95,5 \pm 4,7$	$95,9 \pm 3,3$
modelo 6	Acc	$91,5 \pm 3,5$	$82,8 \pm 9,2$	$88,5 \pm 5,6$	$91,3 \pm 4,3$
	Sens	$81,2 \pm 8,5$	$66,6 \pm 23,4$	$71,6 \pm 7,3$	$89,1 \pm 7,7$
	Esp	$95,0 \pm 2,6$	$88,3 \pm 4,9$	$94,2 \pm 5,8$	$91,9 \pm 5,1$
modelo 7	Acc	$93,6 \pm 4,7$	$87,8 \pm 5,3$	$91,9 \pm 4,4$	$98,0 \pm 2,2$
	Sens	$86,6 \pm 6,5$	$73,0 \pm 14,2$	$83,7 \pm 11,3$	$100,0 \pm 0,0$
	Esp	$96,3 \pm 4,6$	$92,8 \pm 3,7$	$94,6 \pm 3,1$	$97,3 \pm 2,9$

5.3.2 Comitê com Estrutura Otimizada das Camadas Totalmente Conectadas das CNNs

Experimentos foram realizados para analisar o uso do comitê com as redes otimizadas. Com base nos resultados obtidos anteriormente, nessa seção, foi investigado as representações GR e O com as demais abordagens. A Tabela 16 apresenta o resultado

Tabela 15 – Resultados experimentais de acurácia (Acc), sensibilidade (Sens) e especificidade (Esp) para as classes investigadas com o modelos de comitês e estrutura baseada em EfficientNet.

Comitê	Métrica	Saudável	Leve	Moderado	Severa
modelo 1	Acc	96,9 ± 2,3	96,6 ± 2,1	99,3 ± 0,9	99,7 ± 0,7
	Sens	97,3 ± 3,7	90,6 ± 7,6	97,3 ± 3,7	100,0 ± 0,0
	Esp	96,8 ± 2,6	98,6 ± 2,0	100,0 ± 0,0	99,6 ± 1,0
modelo 2	Acc	91,9 ± 1,5	83,4 ± 3,8	86,1 ± 6,2	96,0 ± 3,2
	Sens	76,9 ± 15,8	67,4 ± 9,4	70,1 ± 14,4	100,0 ± 0,0
	Esp	96,9 ± 3,75	88,7 ± 5,0	91,3 ± 6,8	94,6 ± 4,3
modelo 3	Acc	92,5 ± 2,7	81,0 ± 5,7	83,0 ± 5,1	94,6 ± 3,2
	Sens	81,0 ± 9,2	58,1 ± 19,7	66,1 ± 11,1	97,2 ± 3,8
	Esp	96,4 ± 2,1	88,7 ± 2,9	88,7 ± 6,3	93,7 ± 4,5
modelo 4	Acc	93,6 ± 3,6	91,2 ± 4,1	95,6 ± 1,5	95,6 ± 1,9
	Sens	89,2 ± 3,5	79,8 ± 16,2	88,0 ± 7,3	95,8 ± 6,3
	Esp	95,4 ± 4,1	95,1 ± 1,9	98,2 ± 3,0	95,5 ± 3,1
modelo 5	Acc	94,9 ± 5,9	93,6 ± 4,3	96,6 ± 1,7	98,0 ± 0,8
	Sens	93,2 ± 8,2	83,8 ± 14,6	93,3 ± 8,2	95,8 ± 6,3
	Esp	95,4 ± 5,5	96,8 ± 2,0	97,7 ± 2,8	98,6 ± 1,2
modelo 6	Acc	90,8 ± 7,2	84,7 ± 7,0	88,8 ± 4,4	93,6 ± 3,0
	Sens	79,6 ± 13,6	70,6 ± 22,8	73,0 ± 9,2	93,0 ± 8,8
	Esp	94,5 ± 5,1	89,6 ± 3,2	94,1 ± 7,1	93,7 ± 4,5
modelo 7	Acc	94,9 ± 6,0	91,2 ± 4,8	94,9 ± 3,4	98,0 ± 2,2
	Sens	91,8 ± 8,8	80,0 ± 16,3	89,1 ± 3,8	97,1 ± 6,4
	Esp	95,9 ± 5,6	95,0 ± 1,1	96,8 ± 3,8	98,2 ± 2,9

da classificação nessa abordagem. Os melhores resultados obtidos de Acc de 98,0% foram com comitê modelo 1, para estrutura baseada em ResNet-50, e Acc de 97,6%, modelo 1, com a rede EfficientNet. Esses valores são próximos aos dados obtidos somente com a representação GR otimizada (98,6% para EfficientNet e 98% para Resnet-50) (Tabela 6). Assim, observa-se que o uso do comitê em redes otimizadas não apresenta um aprimoramento elevado nos índices de classificação baseado no esforço computacional requerido para implementação. Neste caso, não foram realizadas as avaliações por classe de lesões uma vez que o desempenho foi inferior às abordagens das redes otimizadas sem aplicação de comitês.

Com objetivo de investigar a diferença estatística entre os comitês (otimizado e não otimizado), foi realizado um teste ANOVA de medições repetidas. Cada um dos cinco *folds* foi considerado uma medição repetida. Nesta avaliação, um resultado de p -valor $< 0,001$ foi obtido mostrando que há diferença estatística entre os modelos otimizados e não otimizados. Então, o teste *post-hoc* foi aplicado aos modelos e os resultados são mostrados na Tabela 17. Os resultados obtidos evidenciam que não há diferença significativa entre as abordagens uma vez que para todas as pares (otimizado e não otimizado) o p -valor foi superior a 0,05.

Tabela 16 – Resultados experimentais de Acc com uso de comitê de classificadores e com otimização para as redes investigadas.

EfficientNet						
	<i>Fold 1</i>	<i>Fold 2</i>	<i>Fold 3</i>	<i>Fold 4</i>	<i>Fold 5</i>	<i>Média Fold</i>
modelo 1	97,4	96,5	99,2	96,6	98,3	97,6 $\pm 1,1$
modelo 2	84,4	86,2	87,5	90,0	93,3	88,3 $\pm 3,5$
modelo 3	88,8	87,0	83,3	89,2	91,6	88,0 $\pm 3,1$
modelo 4	95,6	89,6	91,6	93,3	96,7	93,4 $\pm 2,9$
modelo 5	96,5	91,3	95,8	94,2	97,5	95,1 $\pm 2,4$
modelo 6	91,3	83,6	89,2	88,3	95,0	89,5 $\pm 4,2$
modelo 7	97,4	87,8	93,3	90,0	97,5	93,2 $\pm 4,3$

Resnet-50						
	<i>Fold 1</i>	<i>Fold 2</i>	<i>Fold 3</i>	<i>Fold 4</i>	<i>Fold 5</i>	<i>Média Fold</i>
modelo 1	99,1	97,4	99,1	95,8	98,3	98,0 $\pm 1,4$
modelo 2	89,6	87,0	88,8	88,8	91,6	89,0 $\pm 1,7$
modelo 3	93,9	93,1	87,5	89,3	91,6	91,1 $\pm 2,7$
modelo 4	100	94,8	95,0	98,3	98,3	97,3 $\pm 2,3$
modelo 5	98,2	92,2	93,3	90,8	97,5	94,4 $\pm 3,3$
modelo 6	95,6	87,0	88,3	90,8	94,2	91,2 $\pm 3,7$
modelo 7	98,2	93,9	90,8	95,0	95,8	94,8 $\pm 2,7$

Tabela 17 – Resultados dos ρ -valores do teste *post – hoc* para os comitês investigados independente da estrutura da rede.

	Otimizado						
	modelo 1	modelo 2	modelo 3	modelo 4	modelo 5	modelo 6	modelo 7
modelo 1	1,000	<0,001	0,002	0,778	0,557	0,005	0,300
modelo 2	0,002	1,000	1,000	0,039	0,076	1,000	0,174
modelo 3	0,002	1,000	1,000	0,026	0,051	0,997	0,119
modelo 4	0,728	0,045	0,112	1,000	1,000	0,256	1,000
modelo 5	0,732	0,045	0,111	1,000	1,000	0,254	1,000
modelo 6	0,002	1,000	1,000	0,027	0,054	0,998	0,125
modelo 7	0,369	0,135	0,303	0,997	1,000	0,576	1,000

5.4 Uma Visão Geral do Método Proposto com Estado da Arte

A Tabela 18 apresenta uma visão geral dos resultados obtidos neste trabalho em relação a outros estudos de métodos computacionais desenvolvidos para imagens histopatológicas da cavidade oral. Todos os estudos apresentados utilizam imagens de displasia, contudo os estudos de Silva *et al.* (2022), Silva *et al.* (2022) e Silva *et al.* (2024) investigaram a classificação das lesões estratificando o grau de displasia, enquanto que os autores em Neves *et al.* (2023) e Longo *et al.* (2023) realizaram as classificações binárias. É possível observar que apesar das diferentes técnicas de extração de atributos e algoritmos de classificação, o resultados obtidos neste estudo traz contribuições relevantes para o estado da arte da área.

Tabela 18 – Análise da metodologia proposta em relação às abordagens existentes para classificação de imagens orais presentes na literatura.

Estudo	Extração de atributos	Classificador	Classificação	$A_{CC}(\%)$
(SILVA <i>et al.</i> , 2022)	Características de textura e morfológicas	HOP	multi-classe	92,40
(SILVA <i>et al.</i> , 2022)	<i>Deep features</i> com AlexNet and Resnet-50	HOP	multi-classe	98,00
(NEVES <i>et al.</i> , 2023)	ResNet50(<i>Deep features</i>) e LIME	Decorate	binário	97,90
(LONGO <i>et al.</i> , 2023)	Inception-V3 e VGG-19	Comitê Homogêneo	binário	97,97
(SILVA <i>et al.</i> , 2024)	Segmentação de Imagem e Normalização com ResNet50 (<i>Deep features</i>)	RF	multiclasse	94,22
Método Proposto	Representação de Atributos Fractais	CNN Otimizada baseada em Efficient-Net	multiclasse	98,60

Conclusão

Neste trabalho, foi desenvolvido um método computacional utilizando atributos fractais, representação de vetor de características em uma matriz 2D, CNN com transferência de aprendizado para analisar imagens histológicas da cavidade bucal. Em seguida, foi investigado o uso comitê de classificadores baseado na regra da soma e aplicado uma otimização de parâmetros de uma CNN por meio de um algoritmo genético. Essa metodologia foi aplicada a dois modelos CNNs bem conhecidos pelos seus ótimos resultados em análise de imagens, a ResNet-50 e a EfficientNet.

Em um primeiro momento, foi realizado uma classificação utilizando as imagens originais e com representação, sem utilização de técnicas de otimização. Nesta classificação, o método GR se destacou ao alcançar acurácia média de 96,6% com a EfficientNet e 97,6% com a ResNet-50, entre os *folds* analisados. Ao aplicar a otimização do modelo, os resultados de todas as representações sofreram ganhos, atingindo 98,6% com a EfficientNet e 97,6% de acurácia com a ResNet-50, também para representação GR. Ao aplicar o comitê de classificadores, obteve-se resultados promissores alcançado 98,6% de acurácia com a EfficientNet e 98,0% com a ResNet-50. Contudo, ao aplicar o comitê nas redes otimizadas pelo AG, não houve um aumento significativo na acurácia.

Dessa forma, dentre todos os resultados discutidos, foi possível confirmar que uso de atributos fractais aliados as técnicas de representação trazem resultados expressivos no contexto de determinação dos níveis de displasia em imagens da cavidade oral. Além disso, evidencia-se que o uso de técnicas de otimização de parâmetros foram promissores para os modelos quando aplicados as associações atributos fractais e representações 2D avaliados nas arquiteturas de redes CNNs. Um avanço nos resultados ainda foi observado quando as associações, fractal e representação, foram investigadas sobre o comitê de CNNs sem uso de otimização. No entanto, quando os processos de otimização e comitês foram explorados, não foi possível notar relevantes contribuições no desempenho dos modelos pelas métricas investigadas. Com isso, de acordo com o experimentos realizados, não justifica a utilização de ambos os métodos AG e comitê combinados.

A implementação do sistema CAD permitiu classificar com sucesso a imagens de dis-

plasias. Os resultados obtidos foram relevantes e mostram que a metodologia apresentada neste trabalho pode ser utilizada como forma de auxiliar especialistas da área da saúde a estudar e analisar lesões histológicas.

6.1 Principais Contribuições

O método desenvolvido mostrou que é possível combinar estratégias de extração de características baseada em geometria fractal e classificação via CNN com estratégias de otimização para encontrar um modelo relevante ao problema. O método ainda evidencia que o uso de representação de vetores 1D em matrizes 2D pode contribuir significativamente para a classificação de imagens histológicas. Além disso, o estudo demonstra que a aplicação de comitê de classificadores e otimização de parâmetros das camadas totalmente conectadas e de treinamento das redes via algoritmo genético contribuiu na quantificação das lesões baseados no grupo de características extraídas, permitindo melhorar a diferenciação dos graus de displasias. A metodologia proposta proporcionou resultados relevantes comparado com outros trabalhos recentes da literatura que abordam a classificação desse grupo de lesão.

Dessa forma, a metodologia apresentada neste trabalho pode ser utilizada como uma ferramenta para auxiliar especialistas na análise de categorização de displasias. O código fonte deste trabalho está disponível no GitHub, <https://github.com/Carvarafael/Fractal-Reshape-representation>.

6.2 Trabalhos Futuros

Para trabalhos futuros, podem ser abordados outros modelos de otimizar a metodologia apresentada. Para isso, pretende-se estudar novas técnicas de representação bem como o uso de outras estruturas de CNN. Além disso, novos estudos podem ser gerados ao aplicar a metodologia em imagens displásicas de outras lesões histológicas. Também torna-se importante avaliar novas abordagens de algoritmos bioinspirados para as camadas convolucionais das arquiteturas das CNNs.

6.3 Contribuição em Produção Bibliográfica

- ❑ CARVALHO, RAFAEL ; SILVA, ADRIANO ; MARTINS, ALESSANDRO ; CARDOSO, SÉRGIO ; FREIRE, GUILHERME ; R. DE FARIA, PAULO ; LOYOLA, ADRIANO ; TOSTA, THAÍNA ; NEVES, LEANDRO ; Z. DO NASCIMENTO, MARCELO. Oral Dysplasia Classification by Using Fractal Representation Images and Convolutional Neural Networks. In: 19th International Conference on Computer Vision Theory and Applications, 2024, Rome. Proceedings of the 19th Inter-

national Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2024. p. 524;

Referências

- ABBAS, A. K.; FAUSTO, N.; MITCHELL, R. N.; KUMAR, V. **Robbins Patologia Básica**. [S.l.]: Elsevier Brasil, 2008.
- ABBASGHOLIPOUR, M.; OMID, M.; KEYHANI, A.; MOHTASEBI, S. S. Color image segmentation with genetic algorithm in a raisin sorting system based on machine vision in variable conditions. **Expert Systems with Applications**, Elsevier, v. 38, n. 4, p. 3671–3678, 2011.
- ADEL, D. *et al.* Oral epithelial dysplasia computer aided diagnostic approach. In: IEEE. **2018 13th International Conference on Computer Engineering and Systems (ICCES)**. [S.l.], 2018. p. 313–318.
- ADEL, D. *et al.* Oral epithelial dysplasia computer aided diagnostic approach. In: IEEE. **2018 13th International Conference on Computer Engineering and Systems (ICCES)**. [S.l.], 2018. p. 313–318.
- AFONSO, L. C. *et al.* A recurrence plot-based approach for parkinson's disease identification. **Future Generation Computer Systems**, Elsevier, v. 94, p. 282–292, 2019.
- AHMED, T.; SABAB, N. **Classification and understanding of cloud structures via satellite images with EfficientUNet**, *SN Computer Science*, **3**, 99. [S.l.]: DOI, 2022.
- 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**, SciELO Brasil, v. 30, p. 2304–1, 2008.
- AXÉLL, T. Occurrence of leukoplakia and some other oral white lesions among 20 333 adult swedish people. **Community dentistry and oral epidemiology**, Wiley Online Library, v. 15, n. 1, p. 46–51, 1987.
- BAIK, J. *et al.* 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.
- BAISH, J. W.; JAIN, R. K. Fractals and cancer. **Cancer research**, AACR, v. 60, n. 14, p. 3683–3688, 2000.

- BARBOSA, C. E. M. **Algoritmos bio-inspirados para solução de problemas de otimização**. 2017. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2017.
- BENDJILLALI, R. I.; BELADGHAM, M.; MERIT, K.; TALEB-AHMED, A. Illumination-robust face recognition based on deep convolutional neural networks architectures. **Indonesian Journal of Electrical Engineering and Computer Science**, v. 18, n. 2, p. 1015–1027, 2020.
- BIANCHI, A. L. **Entendendo de vez a convolução: base para processamento de imagens**. 2020. Disponível em: <<https://tinyurl.com/4xzn3v73>>. Acesso em: 12/6/2022.
- BIRD, N.; PERRIER, E. Multiscale percolation properties of a fractal pore network. **Geoderma**, Elsevier, v. 160, n. 1, p. 105–110, 2010.
- BRACHMANN, A.; REDIES, C. Defining self-similarity of images using features learned by convolutional neural networks. **Electronic Imaging**, Society for Imaging Science and Technology, v. 2017, n. 14, p. 188–194, 2017.
- BROADBENT, S. R.; HAMMERSLEY, J. M. Percolation processes: I. crystals and mazes. In: CAMBRIDGE UNIVERSITY PRESS. **Mathematical proceedings of the Cambridge philosophical society**. [S.l.], 1957. v. 53, n. 3, p. 629–641.
- CĂLIMAN, A.; IVANOVICI, M. Psoriasis image analysis using color lacunarity. In: IEEE. **2012 13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)**. [S.l.], 2012. p. 1401–1406.
- CAPUTO, L. F. G.; GITIRANA, L.; MANSO, P. Técnicas histológicas. **Molinario EM, Caputo LFG, Amendoeira MRR. Conceitos e métodos para formação de profissionais em laboratórios de saúde**. Rio de Janeiro: Escola Politécnica de Saúde Joaquim Venâncio, p. 138–140, 2010.
- CHAN, H.-P.; HADJIISKI, L. M.; SAMALA, R. K. Computer-aided diagnosis in the era of deep learning. **Medical physics**, Wiley Online Library, v. 47, n. 5, p. e218–e227, 2020.
- COHEN-HADRIA, A.; PEETERS, G. Music structure boundaries estimation using multiple self-similarity matrices as input depth of convolutional neural networks. In: AUDIO ENGINEERING SOCIETY. **Audio Engineering Society Conference: 2017 AES International Conference on Semantic Audio**. [S.l.], 2017.
- DEAN, P. A new monte carlo method for percolation problems on a lattice. In: CAMBRIDGE UNIVERSITY PRESS. **Mathematical Proceedings of the Cambridge Philosophical Society**. [S.l.], 1963. v. 59, n. 2, p. 397–410.
- 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.
- ESTEBSARI, A.; RAJABI, R. Single residential load forecasting using deep learning and image encoding techniques. **Electronics**, MDPI, v. 9, n. 1, p. 68, 2020.

- FARIA, M. P. P.; JULIA, E. S.; FERNANDES, H. C.; NASCIMENTO, M. Z. do; JULIA, R. M. da S. Investigating the performance of optimization techniques on deep learning models to identify dota2 game events. In: **VISIGRAPP (4: VISAPP)**. [S.l.: s.n.], 2023. p. 881–888.
- FARIA, P. R. de *et al.* 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.
- FOOTE, J. Visualizing music and audio using self-similarity. In: **Proceedings of the seventh ACM international conference on Multimedia (Part 1)**. [S.l.: s.n.], 1999. p. 77–80.
- GARTNER, L. P. **Tratado de Histologia Em Cores**. [S.l.]: Elsevier Brasil, 2007.
- GAURAV, S. **Ensemble Methods in Machine Learning: Bagging Versus Boosting**. 2020. Disponível em: <<https://www.pluralsight.com/guides/ensemble-methods:-bagging-versus-boosting>>. Acesso em: 26/7/2023.
- GONÇALVES, C. B.; SOUZA, J. R.; FERNANDES, H. Cnn architecture optimization using bio-inspired algorithms for breast cancer detection in infrared images. **Computers in Biology and Medicine**, Elsevier, v. 142, p. 105205, 2022.
- GONSALVES, P. H. **Geometria Fractal e o Caos**. 2015. Disponível em: <<https://sites.unicentro.br/wp/petfisica/2015/10/16/1311/>>. Acesso em: 20/6/2023.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016. <<http://www.deeplearningbook.org>>.
- GURUCHARAN, M. K. Basic cnn architecture: Explaining 5 layers of convolutional neural network. 2020. Disponível em: <<https://www.upgrad.com/blog/basic-cnn-architecture/>>. Acesso em: 25/5/2021.
- HAYKIN, S. S.; HAYKIN, S. S.; HAYKIN, S. S.; ELEKTROINGENIEUR, K.; HAYKIN, S. S. **Neural networks and learning machines**. [S.l.]: Pearson education Upper Saddle River, 2009. v. 3.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. **CoRR**, abs/1512.03385, 2015. Disponível em: <<http://arxiv.org/abs/1512.03385>>.
- HE, W. *et al.* Progress of machine vision in the detection of cancer cells in histopathology. **IEEE Access**, IEEE, v. 10, p. 46753–46771, 2022.
- 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.
- HUSSAIN, M.; BIRD, J. J.; FARIA, D. R. A study on cnn transfer learning for image classification. In: SPRINGER. **UK Workshop on computational Intelligence**. [S.l.], 2018. p. 191–202.
- HUTTUNEN, M. J. *et al.* Multiphoton microscopy of the dermoepidermal junction and automated identification of dysplastic tissues with deep learning. **Biomedical optics express**, Optical Society of America, v. 11, n. 1, p. 186–199, 2020.

- IVANOVICI, M.; RICHARD, N.; DECEAN, H. Fractal dimension and lacunarity of psoriatic lesions—a colour approach. **medicine**, v. 6, n. 4, p. 7, 2009.
- JOTHI, J. A. A. A survey on automated cancer diagnosis from histopathology images. **Artificial Intelligence Review**, Springer, v. 48, n. 1, p. 31–81, 2017.
- JUNQUEIRA, L. C. U.; CARNEIRO, J. **Histologia básica: texto e atlas**. [S.l.]: Guanabara Koogan, 2013.
- KASSANI, S. H.; KASSANI, P. H.; WESOŁOWSKI, M. J.; SCHNEIDER, K. A.; DETERS, R. Classification of histopathological biopsy images using ensemble of deep learning networks. **arXiv preprint arXiv:1909.11870**, 2019.
- KRAWCZYK, B.; MINKU, L. L.; GAMA, J.; STEFANOWSKI, J.; WOŹNIAK, M. Ensemble learning for data stream analysis: A survey. **Information Fusion**, Elsevier, v. 37, p. 132–156, 2017.
- KRISHNAN, M. M. R. *et al.* Automated classification of cells in sub-epithelial connective tissue of oral sub-mucous fibrosis—an svm based approach. **Computers in biology and medicine**, Elsevier, v. 39, n. 12, p. 1096–1104, 2009.
- KRISHNAN, M. M. R. *et al.* Textural characterization of histopathological images for oral sub-mucous fibrosis detection. **Tissue and Cell**, Elsevier, v. 43, n. 5, p. 318–330, 2011.
- KUMAR, V.; ABBAS, A. K.; ASTER, J. C. **Robbins patologia básica**. 9. ed. [S.l.]: Elsevier Brasil, 2013.
- KUMAR, V.; ASTER, J. C.; ABBAS, A. **Robbins and Cotran Patologia-Bases Patológicas das Doenças**. [S.l.]: Elsevier Brasil, 2010.
- LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, Ieee, v. 86, n. 11, p. 2278–2324, 1998.
- LECUN, Y. A.; BOTTOU, L.; ORR, G. B.; MÜLLER, K.-R. Efficient backprop. In: **Neural networks: Tricks of the trade**. [S.l.]: Springer, 2012. p. 9–48.
- LEI, X.; PAN, H.; HUANG, X. A dilated cnn model for image classification. **IEEE Access**, IEEE, v. 7, p. 124087–124095, 2019.
- LENNON, F. E. *et al.* Lung cancer—a fractal viewpoint. **Nature reviews Clinical oncology**, Nature Publishing Group, v. 12, n. 11, p. 664–675, 2015.
- LONGO, L. H. d. C. *et al.* Classification of multiple h&e images via an ensemble computational scheme. **Entropy**, MDPI, v. 26, n. 1, p. 34, 2023.
- LOPES, R.; BETROUNI, N. Fractal and multifractal analysis: a review. **Medical image analysis**, Elsevier, v. 13, n. 4, p. 634–649, 2009.
- LU, S.; LU, Z.; ZHANG, Y.-D. Pathological brain detection based on alexnet and transfer learning. **Journal of computational science**, Elsevier, v. 30, p. 41–47, 2019.
- LUMINI, A.; NANNI, L. Convolutional neural networks for atc classification. **Current pharmaceutical design**, Bentham Science Publishers, v. 24, n. 34, p. 4007–4012, 2018.

- MALARZ, K. Simple cubic random-site percolation thresholds for neighborhoods containing fourth-nearest neighbors. **Physical Review E**, APS, v. 91, n. 4, p. 043301, 2015.
- MANDELBROT, B. B.; MANDELBROT, B. B. **The fractal geometry of nature**. [S.l.]: WH freeman New York, 1982. v. 1.
- MENDES, R. d. L. *et al.* Abordagens evolutivas para otimização de redes neurais convolucionais baseadas em algoritmos genéticos. Universidade Federal de Uberlândia, 2021.
- MONTERO, P. H.; PATEL, S. G. Cancer of the oral cavity. **Surgical Oncology Clinics**, Elsevier, v. 24, n. 3, p. 491–508, 2015.
- MORRISON, P. **Les Objets Fractals: Forme, Hasard et Dimension**. [S.l.]: JSTOR, 1975.
- MÜLLER, S. Oral epithelial dysplasia, atypical verrucous lesions and oral potentially malignant disorders: focus on histopathology. **Oral surgery, oral medicine, oral pathology and oral radiology**, Elsevier, v. 125, n. 6, p. 591–602, 2018.
- NAJARAN, M. H. T. A genetic programming-based convolutional deep learning algorithm for identifying covid-19 cases via x-ray images. **Artificial Intelligence in Medicine**, Elsevier, v. 142, p. 102571, 2023.
- NANNI, L.; LUMINI, A.; GHIDONI, S. Ensemble of deep learned features for melanoma classification. **arXiv preprint arXiv:1807.08008**, 2018.
- NASCIMENTO, D. S.; CANUTO, A. M.; COELHO, A. L. An empirical analysis of meta-learning for the automatic choice of architecture and components in ensemble systems. In: IEEE. **2014 Brazilian Conference on Intelligent Systems**. [S.l.], 2014. p. 1–6.
- NEVES, L. A. *et al.* Classification of h&e images via cnn models with xai approaches, deepdream representations and multiple classifiers. In: **Proceedings**. [S.l.: s.n.], 2023.
- NEVES, L. A. *et al.* Multi-scale lacunarity as an alternative to quantify and diagnose the behavior of prostate cancer. **Expert systems with applications**, Elsevier, v. 41, n. 11, p. 5017–5029, 2014.
- NIKOLAIDIS, N.; NIKOLAIDIS, I. N.; TSOUROS, C. C. A variation of the box-counting algorithm applied to colour images. **arXiv preprint arXiv:1107.2336**, 2011.
- OMS. **Câncer**. 2023. Disponível em: <https://www.who.int/health-topics/cancer#tab=tab_1>
- OPAS. **Câncer**. 2020. Disponível em: <<https://tinyurl.com/27aw2e7k>>
- PEREIRA, D. C. *et al.* Classification of lymphomas images with polynomial strategy: An application with ridge regularization. In: IEEE. **2022 35th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)**. [S.l.], 2022. v. 1, p. 258–263.
- POLIKAR, R. Ensemble learning. In: **Ensemble machine learning**. [S.l.]: Springer, 2012. p. 1–34.

- REIS, B. M. da F.; DOMINGUES, M. O.; MACAU, E. E. Uso de transformada wavelet discreta e gráfico de recorrência para caracterização do sistema de rössler. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, v. 6, n. 1, 2018.
- RIBEIRO, M. G. *et al.* Analysis of the influence of color normalization in the classification of non-hodgkin lymphoma images. In: IEEE. **2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)**. [S.l.], 2018. p. 369–376.
- ROBERTO, G. F. Associação entre atributos manuais e aprendizado profundo baseada em geometria fractal para classificação de imagens histológicas. Universidade Federal de Uberlândia, 2021.
- ROBERTO, G. F.; NEVES, L. A.; LUMINI, A.; MARTINS, A. S.; NASCIMENTO, M. Z. d. An ensemble of learned features and reshaping of fractal geometry-based descriptors for classification of histological images. **Pattern Analysis and Applications**, Springer, v. 27, n. 1, p. 8, 2024.
- ROBERTO, G. F. *et al.* Features based on the percolation theory for quantification of non-hodgkin lymphomas. **Computers in biology and medicine**, Elsevier, v. 91, p. 135–147, 2017.
- ROBERTO, G. F. *et al.* Detection of covid-19 in chest x-ray images using percolation features and hermite polynomial classification. In: SPRINGER. **Iberoamerican Congress on Pattern Recognition**. [S.l.], 2023. p. 163–177.
- ROSEBROCK, A. **deep learning for computer vision with python**. [S.l.]: PYIMAGESEARCH, 2017. v. 1.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, American Psychological Association, v. 65, n. 6, p. 386, 1958.
- ROY, K.; BANIK, D.; BHATTACHARJEE, D.; NASIPURI, M. Patch-based system for classification of breast histology images using deep learning. **Computerized Medical Imaging and Graphics**, Elsevier, v. 71, p. 90–103, 2019.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **nature**, Nature Publishing Group, v. 323, n. 6088, p. 533–536, 1986.
- SAGI, O.; ROKACH, L. Ensemble learning: A survey. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, Wiley Online Library, v. 8, n. 4, p. e1249, 2018.
- SALAKHUTDINOV, R. R. Learning in markov random fields using tempered transitions. **Advances in neural information processing systems**, v. 22, 2009.
- SANT’ANA, J. M. D. A. *et al.* 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.

SANTOS, A. G. Análise de métodos de otimização de parâmetros e tempo de inferência para modelos de aprendizagem profunda. 2019.

SAVYAKHOSLA. **CNN | Introduction to Pooling Layer**. 2019. Disponível em: <<https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/>>. Acesso em: 19/5/2021.

SCULLY, C. Challenges in predicting which oral mucosal potentially malignant disease will progress to neoplasia. **Oral diseases**, Wiley Online Library, v. 20, n. 1, p. 1–5, 2014.

SILVA, A. B. **Métodos Computacionais para Análise e Classificação de Displasias em Imagens da Cavidade Bucal**. 2019. Tese (dissertation) — Universidade Federal de Uberlândia, 2019.

SILVA, A. B. *et al.* 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.

SILVA, A. B. *et al.* Oralepitheliumdb: A dataset for oral epithelial dysplasia image segmentation and classification. **Journal of Imaging Informatics in Medicine**, Springer, p. 1–20, 2024.

SILVA, A. B. *et al.* Assessment of the association of deep features with a polynomial algorithm for automated oral epithelial dysplasia grading. In: **2022 35th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)**. [S.l.: s.n.], 2022. v. 1, p. 264–269.

SILVA, A. B. *et al.* Métodos computacionais para análise e classificação de displasias em imagens da cavidade bucal. Universidade Federal de Uberlândia, 2019.

SINGH, S. P. **Fully Connected Layer: The brute force layer of a Machine Learning model**. 2018. Disponível em: <<https://iq.opengenus.org/fully-connected-layer/>>. Acesso em: 25/5/2021.

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.

SRIVASTAVA, N.; HINTON, G.; KRIZHEVSKY, A.; SUTSKEVER, I.; SALAKHUTDINOV, R. Dropout: a simple way to prevent neural networks from overfitting. **The journal of machine learning research**, JMLR. org, v. 15, n. 1, p. 1929–1958, 2014.

TAN, C. *et al.* **A Survey on Deep Transfer Learning**. 2018.

TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: PMLR. **International conference on machine learning**. [S.l.], 2019. p. 6105–6114.

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.

- 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.
- TOSTA, T. A. A. *et al.* Método computacional para segmentação não supervisionada de imagens histológicas de linfoma. Universidade Federal de Uberlândia, 2016.
- VASILJEVIC, J. *et al.* Application of multifractal analysis on microscopic images in the classification of metastatic bone disease. **Biomedical microdevices**, Springer, v. 14, n. 3, p. 541–548, 2012.
- WANG, Z.; OATES, T. Encoding time series as images for visual inspection and classification using tiled convolutional neural networks. In: **Workshops at the twenty-ninth AAAI conference on artificial intelligence**. [S.l.: s.n.], 2015.
- WANG, Z.; OATES, T. Imaging time-series to improve classification and imputation. In: **Twenty-Fourth International Joint Conference on Artificial Intelligence**. [S.l.: s.n.], 2015.
- 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.
- WOLK, R.; LINGEN, M. W. Proceedings of the north american society of head and neck pathology companion meeting, new orleans, la, march 12, 2023: Oral cavity dysplasia: Why does histologic grading continue to be contentious? **Head and Neck Pathology**, Springer, p. 1–7, 2023.
- WRIGHT, D.; ADDIS, B. J.; LEONG, A. S. **Diagnostic lymph node pathology**. [S.l.]: CRC Press, 2011.
- 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.
- YANASE, J.; TRIANTAPHYLLOU, E. A systematic survey of computer-aided diagnosis in medicine: Past and present developments. **Expert Systems with Applications**, Elsevier, v. 138, p. 112821, 2019.
- YOSINSKI, J.; CLUNE, J.; BENGIO, Y.; LIPSON, H. How transferable are features in deep neural networks? **arXiv preprint arXiv:1411.1792**, 2014.
- ZHOU, Z.-H. **Ensemble methods: foundations and algorithms**. [S.l.]: CRC press, 2012.
- ZIFF, R. M. Test of scaling exponents for percolation-cluster perimeters. **Physical review letters**, APS, v. 56, n. 6, p. 545, 1986.

