



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de São José do Rio Preto

Jaqueline Junko Tenguam

Ensemble Learning como Estratégia para
Investigar Imagens H&E utilizando
Duplo Estágio de Seleção de Atributos

São José do Rio Preto
2023

Jaqueline Junko Tengum

Ensemble Learning como Estratégia para
Investigar Imagens H&E utilizando
Duplo Estágio de Seleção de Atributos

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação, junto ao Programa de Pós-Graduação PPGCC – Programa de Pós-Graduação em Ciência da Computação, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

Financiadora: CAPES

Orientador: Prof. Dr. Leandro Alves
Neves

São José do Rio Preto
2023

T292e

Tengum, Jaqueline Junko

Ensemble learning como estratégia para investigar imagens H&E utilizando duplo estágio de seleção de atributos / Jaqueline Junko Tengum. -- São José do Rio Preto, 2023

140 f. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto

Orientador: Leandro Alves Neves

1. Ensemble learning. 2. Descritores handcrafted. 3. Descritores deep learned. 4. Imagens histológicas. 5. Duplo estágio de seleção de atributos. I. Título.

Jaqueline Junko Tengum

Ensemble Learning como Estratégia para
Investigar Imagens H&E utilizando
Duplo Estágio de Seleção de Atributos

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação, junto ao Programa de Pós-Graduação PPGCC – Programa de Pós-Graduação em Ciência da Computação, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

Financiadora: CAPES

Comissão Examinadora

Prof. Dr. Leandro Alves Neves
UNESP – Câmpus de São José do Rio Preto
Orientador

Prof. Dr. Denis Henrique Pinheiro Salvadeo
UNESP – Câmpus de Rio Claro

Prof. Dr. André Ricardo Backes
UFSCar – Universidade Federal de São Carlos

São José do Rio Preto
01 de fevereiro de 2023

Ao meu querido amigo Rodolfo Rodrigues Simões (in memoriam).

Agradecimentos

Agradeço ao meu orientador Prof. Dr. Leandro Alves Neves por todo o apoio, compreensão e incentivo que demonstrou durante todo o tempo em que estivemos trabalhando juntos.

Agradeço aos meus amigos e familiares por sempre estarem ao meu lado, apoiando e incentivando a minha carreira acadêmica e, acima de tudo, por sempre se orgulharem de tudo o que conquistei até aqui.

O orientador deste projeto agradece o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), #313643/2021-0, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), #2022/03020-1, e da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), #APQ01129-21 e #APQ-00578-18.

Finalmente, agradeço ao apoio financeiro fornecido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), código de financiamento 001, para o desenvolvimento deste trabalho.

"Continue a nadar."

- Dory (Procurando Nemo)

Resumo

Neste trabalho, é apresentada uma investigação de *ensemble learning* para o contexto de imagens médicas, especificamente o reconhecimento de padrões em amostras histológicas tingidas com Hematoxilina e Eosina. As bases exploradas foram representativas do câncer colorretal, displasia oral epitelial, linfomas não-Hodgkin e tecidos hepáticos. A estratégia de *ensemble learning* foi considerada a partir de múltiplos descritores, tais como *deep learned* e *handcrafted*, e múltiplos classificadores. Os descritores *deep learned* foram calculados explorando cinco distintas arquiteturas de redes neurais convolucionais. Os descritores *handcrafted* foram representativos das categorias fractais multidimensionais e multiescala, Haralick e *Local Binary Pattern*. As principais combinações de descritores, consequentemente de técnicas, foram obtidas por meio de duplo estágio de seleção de atributos (ranqueamento com meta-heurísticas) e classificadas via um *ensemble* de classificadores compostos por múltiplos classificadores: *Support Vector Machine*, *Naive Bayes*, *Random Forest*, Regressão Logística e *K-Nearest Neighbors*. As taxas de acurácias foram valores entre 90,72% e 100,00%, com alguns destaques envolvendo combinações de técnicas capazes de aprimorar as acurácias em diferentes contextos de imagens histológicas; análise dos atributos presentes nas melhores soluções obtidas; e definição de combinações de *ensembles* com desempenhos competitivos em relação aos disponíveis na literatura, explorando um número reduzido de atributos.

Palavras-chave: *Ensemble learning*. Descritores *handcrafted*. Descritores *deep learned*. Imagens histológicas. Duplo estágio de seleção de atributos.

Abstract

In this work, an ensemble learning investigation is presented for the context of medical images, specifically the recognition of patterns in histological samples stained with Hematoxylin and Eosin. The datasets explored were representative of colorectal cancer, oral epithelial dysplasia, non-Hodgkin lymphoma, and liver tissue. The ensemble learning strategy was considered from multiple descriptors, such as deep learned and handcrafted, and multiple classifiers. The deep learned features were calculated by exploring five different architectures of convolutional neural network. The handcrafted features were representative of the multidimensional and multiscale fractal, Haralick and Local Binary Pattern categories. The main combinations of descriptors, consequently of techniques, were obtained through a two-stage feature selection (ranking with metaheuristics) and classified via an ensemble of classifiers composed of multiple classifiers: Support Vector Machine, Naive Bayes, Random Forest, Logistic Regression, and K-Nearest Neighbors. The accuracy rates were values between 90.72% and 100.00%, with some highlights involving combinations of techniques capable of improving accuracy in different histological image contexts; analysis of the features present in the best solutions obtained; and definition of ensemble combinations with competitive performances compared to those available in the literature, exploring a reduced number of descriptors.

Keywords: *Ensemble learning. Handcrafted features. Deep learned features. Histological images. Two-stage feature selection.*

Lista de Figuras

2.1	Ângulos θ considerados para o cálculo da matriz de co-ocorrência. . .	31
2.2	Exemplificação do processo de construção da matriz de co-ocorrência, com $\delta = 1$ e $\theta = 0^\circ$: (a) Imagem original, com a indicação de pares de <i>pixels</i> vizinhos; (b) Estágio de incrementação do valor na matriz após analisar os <i>pixels</i> vizinhos destacados em (a); (c) Matriz de co-ocorrência resultante da imagem (a).	32
2.3	Diferentes configurações de vizinhança (PN, R) para o método LBP. .	33
2.4	Obtenção do código LBP, considerando $R = 1$ e $PN = 8$	33
2.5	Exemplo de arquitetura CNN, considerando um problema de classificação de veículos.	36
2.6	Exemplo do processo de convolução de uma imagem 4×4 com um filtro de tamanho 2×2 e o resultado da operação.	37
2.7	Saídas de cada camada (horizontal) de uma CNN. Cada imagem é um mapa de características correspondente à saída das características aprendidas na camada anterior, detectada em cada posição da imagem.	38
2.8	Operação de <i>max pooling</i>	39
2.9	Ilustrações de funções de ativação: Sigmoides, à esquerda; ReLU, à direita.	40
2.10	Arquitetura do modelo VGG-19.	43
2.11	Ilustração de um bloco residual na arquitetura ResNet-50.	43
2.12	Arquitetura ResNet-50.	44
2.13	Arquitetura da rede Inception v3.	45

2.14	Redução do tamanho do <i>grid</i> . O diagrama na direita representa a mesma solução à esquerda, mas da perspectiva de tamanho do <i>grid</i> ao invés de operações.	46
2.15	Arquitetura DenseNet-121.	47
2.16	Arquitetura EfficientNet-B0.	48
2.17	Duas principais arquiteturas de <i>ensemble</i> . S_i , h_i e $H(x)$ indicam a i -ésima amostra de dados, o i -ésimo classificador e a regra de combinação, respectivamente.	56
2.18	Ilustração das quatro abordagens para a criação de um <i>ensemble</i>	57
2.19	Combinação baseada em regras fixas em um problema com quatro classes e um <i>ensemble</i> de três classificadores.	59
2.20	Gráfico de comparação entre três curvas ROC.	61
2.21	Exemplos de imagens histológicas H&E.	63
4.1	Ilustração do método proposto.	88
4.2	Ilustração do método <i>gliding box</i> considerando $L = 3$ (a) e $L = 5$ (b), com um hipercubo, representado em vermelho, deslocando-se na imagem até o canto inferior.	90
4.3	Esquema do <i>ensemble</i> proposto em que as caixas com linhas tracejadas definem que apenas um dos seus componentes foi utilizado para compor o <i>ensemble</i>	104
4.4	Exemplos de imagens NHL (a), OED (b), CR (c), LG (d) e LA (e). . . .	105
5.1	Ocorrências de descritores <i>deep learned</i> presentes nas <i>top</i> 10 soluções das bases CR, LG e NHL.	112
5.2	Ocorrência dos descritores <i>handcrafted</i> nas <i>top</i> 10 soluções para as bases LA, OED, NHL e LG.	113

Lista de Tabelas

2.1	Matriz de confusão com as quatro classificações possíveis.	60
3.1	Sumarização dos principais trabalhos correlatos que utilizaram <i>ensemble</i> de descritores para análises de imagens médicas.	86
3.2	Visão geral dos trabalhos que utilizaram duplo estágio de seleção em diferentes contextos de análise.	87
4.1	Ilustração de uma matriz de probabilidades.	91
4.2	Resumo de cada modelo CNN investigado neste projeto.	100
4.3	Nome, medida de distância considerada e o total de características <i>handcrafted</i> resultante.	102
4.4	Número de descritores extraídos e sua referente camada em cada arquitetura.	102
4.5	Informações gerais das bases de imagens utilizadas no trabalho. . . .	106
5.1	Acurácias médias obtidas pelo método proposto, utilizando dois estágios de seleção (ReliefF + <i>Wrapper</i>). As melhores acurácias estão destacadas em negrito.	109
5.2	Principais combinações em cada conjunto H&E a partir do critério maior acurácia com o menor número de descritores.	110
5.3	Acurácias fornecidas pelas arquiteturas CNN após processo de <i>fine-tuning</i> : os melhores resultados estão em negrito.	114
5.4	Teste de Friedman: ranqueamento dos métodos.	116

5.5	Visão geral dos valores de acurácias obtidos por diferentes abordagens para a classificação de imagens de câncer colorretal (CR).	117
5.6	Visão geral dos valores de acurácias obtidos por diferentes abordagens para a classificação de imagens de displasia oral (OED).	118
5.7	Visão geral dos valores de acurácias obtidos por diferentes abordagens para a classificação de imagens de fígado (LA e LG).	118
5.8	Visão geral dos valores de acurácias obtidos por diferentes abordagens para a classificação de imagens de linfomas (NHL).	118
A.1	Acurácias e outras informações complementares obtidas nos 10 testes executados para cada combinação <i>ensemble</i> de descritores + seletor <i>wrapper</i> no <i>dataset</i> CR.	138
A.2	Acurácias e outras informações complementares obtidas nos 10 testes executados para cada combinação <i>ensemble</i> de descritores + seletor <i>wrapper</i> no <i>dataset</i> OED.	139
A.3	Acurácias e outras informações complementares obtidas nos 10 testes executados para cada combinação <i>ensemble</i> de descritores + seletor <i>wrapper</i> no <i>dataset</i> LA.	139
A.4	Acurácias e outras informações complementares obtidas nos 10 testes executados para cada combinação <i>ensemble</i> de descritores + seletor <i>wrapper</i> no <i>dataset</i> LG.	140
A.5	Acurácias e outras informações complementares obtidas nos 10 testes executados para cada combinação <i>ensemble</i> de descritores + seletor <i>wrapper</i> no <i>dataset</i> NHL.	140

Lista de Abreviações

ACC Acurácia

AHP *Adaptive Hybrid Pattern*

AUC *Area Under the ROC Curve*

bGWO binary Grey Wolf Optimization

CLBP *Completed LBP*

CLM *CodebookLess Model*

CNN *Convolutional Neural Network*

COL *Color descriptor*

CR Colorretal

DL *Deep Learned*

FBSIF *Full Binarized Statistical Image Feature*

GA Algoritmo Genético

GLCM *Gray-level Co-occurrence Matrix*

GOLD *Gaussian Of Local Descriptors*

H&E Hematxilina-Eosina

HF *Handcrafted*

HOG *Histogram of Oriented Gradients*

HSL *Hue Saturation Luminosity*

ILSVRC *ImageNet Large Scale Visual Recognition Challenge*

ISIC *International Skin Imaging Collaboration*

KNN *K-Nearest Neighbors*

LA *Liver Aging*

LBP *Local Binary Patterns*

LET *LETRIST (Locally Encoded Transform Feature Histogram) Descriptor*

LG *Liver Gender*

LightGBM *Light Gradient-Boosting Machine*

LTP *Local Ternary Patterns*

MLPQ *Multithreshold LPQ*

MOR *Morphological Descriptor*

MRI *Magnetic Resonance Imaging*

NHL *Linfoma Não-Hodgkin*

OED *Displasia Epitelial Oral*

PSO *Particle Swarm Optimization*

RBF *Radial Basis Function*

RGB *Red Green Blue*

RIC *Rotation Invariant Co-occurrence*

ROC *Receiver Operating Characteristic*

ROI *Region of Interest*

SIFT *Scale-invariant Feature Transform*

SVM *Support Vector Machine*

Sumário

1	INTRODUÇÃO	19
1.1	Motivação e Justificativas	23
1.2	Objetivos	24
1.3	Considerações Finais sobre a Introdução	25
2	FUNDAMENTAÇÃO TEÓRICA	27
2.1	Extração de Características	27
2.2	Técnicas Fractais	28
2.2.1	Dimensão Fractal	29
2.2.2	Lacunaridade	29
2.2.3	Percolação	30
2.3	Descritores de Haralick	30
2.4	<i>Local Binary Patterns</i> – LBP	32
2.5	<i>Deep Learning</i>	34
2.6	Redes Neurais Convolucionais	35
2.6.1	Camada Convolucional	36
2.6.2	<i>Pooling</i>	38
2.6.3	Função de Ativação	39
2.6.4	Transferência de Aprendizado	40
2.7	Arquiteturas de Redes Neurais Convolucionais	41
2.7.1	VGG-19	42
2.7.2	ResNet-50	42

2.7.3	Inception v3	44
2.7.4	DenseNet-121	46
2.7.5	EfficientNet	47
2.8	Técnicas para Seleção de Características	49
2.9	Classificadores	52
2.10	<i>Ensemble Learning</i>	55
2.11	Métricas para Avaliação de Desempenho	58
2.12	Imagens Histológicas	62
2.13	Considerações Finais sobre a Fundamentação Teórica	62
3	TRABALHOS RELACIONADOS	65
3.1	Técnicas Fractais Multidimensionais e Multiescala	65
3.2	Descritores de Haralick	68
3.3	Descritores LBP	70
3.4	Redes Neurais Convolucionais	71
3.5	<i>Ensemble</i> de Descritores <i>Deep Learned</i> e <i>Handcrafted</i> sem o uso <i>Ensemble</i> de Classificadores	75
3.6	<i>Ensemble</i> de Descritores <i>Deep Learned</i> e <i>Handcrafted</i> com o uso de <i>Ensemble</i> de Classificadores	78
3.7	Duplo estágio de Seleção de Atributos	82
3.8	Considerações sobre os Trabalhos Relacionados	84
4	MATERIAIS E MÉTODOS	88
4.1	Descritores <i>handcrafted</i> – Técnicas fractais multidimensionais e multiescala	89
4.1.1	Dimensão fractal baseada em abordagem probabilística	89
4.1.2	Dimensão fractal baseada em abordagem de fusão de caixas	91
4.1.3	Dimensão fractal de Higuchi multidimensional e multiescala	92

4.1.4	Lacunaridade	94
4.1.5	Percolação multidimensional e multiescala	95
4.1.6	Métricas obtidas das Curvas dos descritores	96
4.2	Descritores <i>handcrafted</i> – Haralick	98
4.3	Descritores <i>handcrafted</i> – <i>Local Binary Pattern</i> (LBP)	99
4.4	Descritores <i>deep learned</i>	99
4.5	<i>Ensemble</i> de descritores	100
4.6	Análises de descritores via duplo estágio de seleção e <i>ensemble</i> de classificadores	101
4.6.1	<i>Ensemble Learning</i>	103
4.7	Contextos de aplicação	104
4.8	Pacotes de <i>software</i> e ambiente para os experimentos	106
4.9	Considerações Finais sobre Materiais e Métodos	107
5	RESULTADOS E DISCUSSÃO	108
5.1	Ocorrências dos descritores nas principais soluções: uma visão geral	111
5.2	Visão Geral frente a diferentes abordagens	114
5.2.1	Observações em relação aos trabalhos correlatos	116
6	CONCLUSÃO	119
6.1	Trabalhos Futuros	121
6.2	Contribuições	121
	REFERÊNCIAS	122
A	RESULTADOS OBTIDOS	138

Capítulo 1

Introdução

Técnicas computacionais para a análise de imagens têm contribuído significativamente para o desenvolvimento e aprimoramento de sistemas de apoio ao diagnóstico de doenças, os denominados *Computer-Aided Detection* (CAD) (ARAÚJO et al., 2017; CHAN; HADJIISKI; SAMALA, 2020; TANG et al., 2009). Para tanto, diversas combinações de métodos são explorados nestes sistemas, sendo alguns deles direcionados para apoiar a detecção de uma doença via análise histopatológica (ARAÚJO et al., 2017; IFTIKHAR; HASSAN; ALQUHAYZ, 2016; JØRGENSEN et al., 2017; KLONOWSKI et al., 2013; RIBEIRO et al., 2019). Este tipo de análise consiste em utilizar microscopia de luz para investigar amostras de tecidos (JUNQUEIRA; CARNEIRO, 2013; NAYAK; MISHRA, 2016). As várias etapas necessárias no processo de preparação das amostras podem interferir na organização das estruturas, fato que proporciona ainda mais dificuldades para um diagnóstico preciso dos diferentes níveis de uma doença sob análise. Somado a tal problemática, este tipo de análise demanda tempo e pode ser suscetível a erros de interpretação, decorrentes de uma variedade de situações envolvendo o histopatologista (AKBAR; GOPI; BABU, 2015; ALTUNBAY et al., 2010; IFTIKHAR; HASSAN; ALQUHAYZ, 2016).

Neste contexto, dois principais grupos de descritores são comumente explorados para o desenvolvimento dos sistemas CAD. O primeiro deles é o de descritores *hand-*

crafted (HF), os quais são definidos por métodos de extração distintos, geralmente visando superar problemas específicos (HASSAN et al., 2022; NANNI; GHIDONI; BRAHNAM, 2017; NANNI et al., 2020; SETHY; BEHERA, 2022). Dentre os atributos *handcrafted*, é possível destacar as abordagens baseadas em técnicas fractais multiescalas e/ou multidimensionais (BAISH; JAIN, 2000; IVANOVICI; RICHARD, 2011; LI et al., 2012; LOPES; BETROUNI, 2009; ROBERTO et al., 2021), Haralick (HARALICK; SHANMUGAM; DINSTEN, 1973) e *Local Binary Pattern* (LBP) (OJALA; PIETIKÄINEN; HARWOOD, 1996). Por exemplo, Haralick e LBP têm sido explorados em diversos contextos de imagens (ALMARAZ-DAMIAN et al., 2020; NKETIAH et al., 2017; WIBMER et al., 2015), com destaque para subsidiar as identificações de subtipos de câncer de pulmão (WANG; YU, 2013), malignidade de amostras de tecido mamário (FONDÓN et al., 2018; HASSAN et al., 2022) e classificação de câncer colorretal (AKBAR; GOPI; BABU, 2015). De maneira complementar, as técnicas fractais multiescalas e/ou multidimensionais também indicaram resultados relevantes para quantificar arquiteturas patológicas de tumores (BAISH; JAIN, 2000; LI et al., 2012; LOPES; BETROUNI, 2009), com destaque para o câncer de próstata (NEVES et al., 2014), linfomas (ROBERTO et al., 2017), neoplasia intraepitelial anal (KLONOWSKI et al., 2018), tumores de mama (ROBERTO et al., 2019), câncer colorretal (RIBEIRO et al., 2019) e lesões psoriáticas (IVANOVICI; RICHARD; DECEAN, 2009).

O segundo grupo é constituído por atributos aprendidos (*learned features*) via redes neurais convolucionais (CNN) (BUDNIK et al., 2017), especialmente. Assim, por exemplo, os modelos de redes são aplicados em amostras de imagens de uma coleção e os parâmetros obtidos durante a fase de treinamento possibilitam a extração de vetores de características (BUDNIK et al., 2017). Este tipo de estratégia tem embasado inúmeras técnicas para analisar imagens médicas (ELMANNAI; HAMDI; ALGARNI, 2021; GREENSPAN; GINNEKEN; SUMMERS, 2016; JASTI et al., 2022;

KASSANI et al., 2019; LITJENS et al., 2017). Isso é possível porque o aprendizado profundo permite que redes CNN, compostas por várias camadas de processamento simples e não-lineares, aprendam representações de dados com vários níveis de abstração. Com isso, as entradas brutas são transformadas em representações abstratas de nível mais alto (LECUN; BENGIO; HINTON, 2015), com aplicações nas mais diferentes tarefas (AL-MASNI et al., 2017; ARAÚJO et al., 2017; KASSANI et al., 2019). É importante destacar que os modelos mais explorados são aqueles que se destacaram com as melhores taxas de acurácias na competição de classificação do banco de imagens ImageNet (RUSSAKOVSKY et al., 2015), tais como: AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), GoogleNet (SZEGEDY et al., 2015a), VGGNet (SIMONYAN; ZISSERMAN, 2014), ResNet (HE et al., 2016), Inception (SZEGEDY et al., 2015a) e outras variantes (ALMARAZ-DAMIAN et al., 2020; NANNI et al., 2020; RAJINIKANTH et al., 2020).

Apesar dos avanços conquistados explorando os grupos de características apontados previamente, pesquisas foram desenvolvidas com o intuito de explorar o desempenho de modelos que combinaram ambos os tipos de atributos, *handcrafted* e *deep learned* (DL) (ALMARAZ-DAMIAN et al., 2020; HASSAN et al., 2022; NANNI; GHIDONI; BRAHNAM, 2018; RAJINIKANTH et al., 2020; SETHY; BEHERA, 2022). Este tipo de combinação é nomeado como *ensemble learning* (HAGERTY et al., 2019; NANNI et al., 2020) e pode envolver também múltiplos algoritmos de classificação. Com isso, é possível ter uma decisão final única, mesmo considerando classificadores independentes, e mais precisa (NANNI et al., 2020; PONTI JR., 2011). O uso de *ensemble* de classificadores treinados com os atributos *handcrafted* e *deep learned* permitiu, por exemplo, concluir que os descritores *handcrafted* melhoraram o desempenho das classificações de células cervicais (NANNI et al., 2020). Outro destaque é o modelo apresentado por Hagerty et al. (2019), em que os autores realizaram um estudo comparativo entre o classificador (regressão logística) treinado somente com as carac-

terísticas *deep learned*, *ensemble* de atributos *handcrafted* e *ensemble* com todos os descritores. Os autores concluíram que o *ensemble* envolvendo todas as características indicaram as melhores taxas de distinções. No escopo de técnicas fractais e modelos de CNN, os autores Roberto et al. (2021) propuseram um *ensemble* entre duas arquiteturas CNN, uma treinada com imagens histológicas e a outra treinada com imagens artificiais, geradas a partir de características obtidas por meio de técnicas fractais. Os autores concluíram que a proposta superou as classificações obtidas pelos algoritmos de classificação e modelos CNN aplicados isoladamente.

Por fim, ainda é possível aplicar ao modelo de *ensemble learning* um estágio de seleção de atributos, com o intuito de reduzir o espaço de busca e aumentar a precisão do sistema, podendo ser realizado via seleção em uma fase (HASSAN et al., 2022; RIBEIRO et al., 2019), com a aplicação de um único seletor ou ainda, com a aplicação de duplo estágio de seleção, como abordado pelos autores (BHOWAL; SEN; SARKAR, 2021; CANDELERO et al., 2020; LI et al., 2022; ZHANG et al., 2018). A dupla seleção de atributos permite conhecer as principais composições para o desenvolvimento de sistemas CAD, tornando o conhecimento mais compreensível para os especialistas.

Esse trabalho está dividido em seis capítulos. No Capítulo 2 são apresentados os conceitos fundamentais para o entendimento do trabalho proposto. O Capítulo 3 aborda alguns dos trabalhos presentes na literatura, relacionados tanto com o tema abordado quanto com as técnicas propostas. No Capítulo 4 é introduzido a metodologia proposta e no Capítulo 5 é apresentado os resultados da aplicação do método proposto e uma discussão. Finalmente, no Capítulo 6 é dada uma breve conclusão do trabalho, algumas das contribuições e trabalhos futuros. O Apêndice A é composto por alguns resultados complementares.

1.1 Motivação e Justificativas

O estudo de diferentes tipos de tecidos presentes nas imagens histológicas, usualmente corados com materiais específicos, tais como Hematoxilina e Eosina (H&E), constituem parte fundamental do processo de análise histopatológica (KATHER et al., 2016). Doenças como o câncer, por exemplo, responsáveis pela segunda principal causa de morte nos Estados Unidos (SIEGEL et al., 2021), são diagnosticados mediante avaliações visuais de amostras de tecidos, por meio da microscopia de luz. Muitas vezes, a detecção precoce da doença possibilita tratamentos menos invasivos e aumenta a possibilidade de cura e/ou sobrevida do paciente. Este processo implica em problemas de interpretação, principalmente no que se refere aos quesitos humanos, tais como a subjetividade e a fadiga, por exemplo. Por outro lado, os sistemas CAD têm um papel fundamental neste processo, visto que podem auxiliar os especialistas com uma segunda opinião. Mais ainda, as diferentes abordagens de aprendizado profundo e arquiteturas CNN têm contribuído significativamente com o desenvolvimento de sistemas CAD mais precisos (GREENSPAN; GINNEKEN; SUMMERS, 2016; LITJENS et al., 2017), inclusive para análises de imagens histológicas (ARAÚJO et al., 2017; NANNI et al., 2020; ROBERTO et al., 2021; SHARMA; KUMAR, 2022; WANG et al., 2022).

Neste cenário, observa-se a versatilidade da estratégia de *ensemble learning*, uma vez que é possível realizar diversas combinações entre características e classificadores para investigar algumas categorias de imagens histológicas (HAGERTY et al., 2019; LI et al., 2018; NANNI; GHIDONI; BRAHNAM, 2018; NANNI et al., 2020; ROBERTO et al., 2021). Nestes trabalhos, os autores exploraram alguns *ensembles* de descritores *handcrafted* e *deep learned* (ALMARAZ-DAMIAN et al., 2020; KUMAR et al., 2022; NANNI; GHIDONI; BRAHNAM, 2018; PONTI JR., 2011; RAJINKANTH et al., 2020; SETHY; BEHERA, 2022), mas não como os propostos e observados aqui, visando conhecer possíveis padrões de combinações de técnicas em

diferentes tipos de amostras histológicas H&E. Alguns exemplos de estruturas *ensemble learning* que ainda podem ser investigadas são: descritores *handcrafted* com um *ensemble* de classificadores; descritores *deep learned* com um *ensemble* de classificadores; descritores *handcrafted* e *deep learned* com um *ensemble* de classificadores. A melhor estrutura pode ser comparada via classificações somente com o uso dos modelos CNN, fato que permitiu responder à pertinência ou não do uso de *ensemble learning* para classificar e reconhecer padrões em imagens representativas do câncer colorretal, displasia oral, linfomas não-Hodgkin e tecidos hepáticos.

Esses *ensembles* ainda podem ser explorados via duplo estágio de seleção. O intuito dessa abordagem é conhecer os descritores mais importantes para compor os sistemas CAD, além de reduzir significativamente um espaço de busca de alta dimensionalidade, como o explorado nesta proposta. Isso permite minimizar problemas comuns em modelos de aprendizagem, tais como situações de *overfitting* ou desempenhos abaixo do esperado (ZHAO et al., 2015). Assim, a intenção é aproveitar as vantagens de cada categoria de seletor, de forma que um esquema de seleção mais robusto resulte em soluções computacionais capazes de fornecer os melhores desempenhos em diferentes contextos de imagens histológicas H&E, com informações envolvendo as principais combinações de descritores e métodos de seleção.

1.2 Objetivos

Neste projeto é apresentada uma proposta de investigação de *ensemble learning*, envolvendo múltiplos descritores, *deep learned* e *handcrafted*, e múltiplos classificadores. A intenção é contribuir com os processos de interpretação e reconhecimento de padrões presentes em imagens com relevância científica. Para tanto, os objetivos específicos foram:

- Quantificar as imagens histológicas a partir de técnicas *handcrafted*, tais como

técnicas fractais multidimensionais (dimensão fractal de Higuchi, dimensão fractal probabilística, dimensão fractal por fusão de caixas, lacunaridade e percolação), descritores de Haralick e LBP, bem como técnicas *deep learned*, aplicando diferentes redes;

- Investigar as principais combinações de descritores a partir de esquemas de duplos estágios de seleção (ranqueamento com meta-heurísticas), a fim de identificar possíveis estruturas capazes de maximizar as taxas de classificações e evitar a presença de *overfitting*;
- Aplicar um *ensemble* de classificadores que permita fornecer taxas de distinções relevantes e sem *overfitting*;
- Conhecer possíveis padrões de *ensemble learning* para o estudo e reconhecimento de padrões em imagens histológicas representativas do câncer colorretal, displasia oral, linfomas não-Hodgkin e tecidos hepáticos.

1.3 Considerações Finais sobre a Introdução

Neste capítulo são abordados alguns dos conceitos iniciais e uma contextualização que envolve o cenário de aplicação da proposta aqui definida. Vimos que na literatura são encontrados diferentes abordagens utilizando a proposta de *ensemble learning* para a análise de diferentes tipos de imagens médicas e como isso impacta nos sistemas CAD. Dentre as estratégias, estão aquelas que utilizam os descritores mais clássicos, como os descritores *handcrafted* LBP e Haralick, ou mesmo descritores mais robustos, como os extraídos das arquiteturas CNN. Neste trabalho será apresentada uma abordagem diferente das encontradas na literatura, com o intuito de investigar as diferentes combinações de *ensembles*, explorando cinco arquiteturas de redes neurais convolucionais como extratores das *deep learned features* e descritores *handcrafted* baseados em texturas, avaliados sob duplo estágio de seleção de atributos e com um *ensemble* de

cinco diferentes classificadores para a investigação e análise de imagens histológicas H&E de relevância científica.

Com isso, buscamos responder questões como: qual(is) a(s) melhor(es) estrutura(s) de *ensemble learning* para a classificação de imagens histológicas no contexto em análise? A estratégia de *ensemble learning* proposta consegue superar as classificações obtidas somente com as arquiteturas CNNs treinadas exclusivamente para essa tarefa? O uso de *ensemble learning* para classificar e reconhecer padrões em imagens representativas de câncer colorretal, linfomas não-hodgkin, displasia oral epitelial e tecidos hepáticos é pertinente ou não?

Capítulo 2

Fundamentação Teórica

Neste capítulo são apresentados os conceitos fundamentais utilizados no trabalho proposto e que serviram para o entendimento do mesmo. Para tanto, na Seção 2.1 é abordado o tema de extração de características e, nas Seções seguintes, 2.2, 2.3 e 2.4, são descritas técnicas que exemplificam a categoria *handcrafted features*. Nas Seções de 2.5 a 2.7 são abordados os temas que envolvem as características obtidas por meio de algoritmos *deep learning*. Neste sentido, são especificados os assuntos *deep learning*, redes neurais convolucionais e as arquiteturas de redes neurais. Técnicas para seleção de características e classificação dos descritores são abordados nas Seções 2.8 e 2.9, respectivamente. A Seção 2.10 é direcionada para apresentar os principais conceitos de *ensemble learning*. Os fundamentos de métricas comumente utilizadas para avaliar os classificadores estão na Seção 2.11. Por fim, informações gerais dos contextos de aplicação da presente proposta estão na Seção 2.12.

2.1 Extração de Características

A área de reconhecimento de padrões envolve a descrição ou a classificação de medidas. Não é uma técnica única, mas sim um amplo conjunto de conhecimentos e técnicas. Uma dessas técnicas consiste na aplicação de algoritmos para extração de ca-

racterísticas nos dados de entrada (WAH, 2008). Uma característica pode ser definida como uma informação quantitativa como, por exemplo, uma medida de comprimento. Inúmeros tipos de características podem ser extraídas, sendo elas de dimensão, cor, forma, geometria e outros (PEDRINI; SCHWARTZ, 2008). Exemplos de técnicas são as abordagens fractais, descritores de Haralick, LBP e características aprendidas, extraídas de arquiteturas de redes neurais convolucionais.

2.2 Técnicas Fractais

A geometria fractal, definida por Mandelbrot (MANDELBROT, 1977), foi introduzida para permitir o estudo e a descrição de formas irregulares ou fragmentadas de elementos da natureza, bem como objetos complexos que a geometria euclidiana não conseguia analisar (LOPES; BETROUNI, 2009). O termo fractal pode ser definido como um objeto geométrico que apresenta propriedade de autossimilaridade, em que uma parte do objeto é idêntico ao todo. A geometria fractal é um modelo matemático capaz de analisar e representar sistemas complexos, os quais não possuem escala de comprimento característico. Um exemplo disso é o formato de um litoral. Quando duas imagens de um litoral são dadas em duas escalas distintas, com 1 cm correspondendo, por exemplo, a 0,1 km ou a 10 km, não é possível afirmar qual escala pertence a qual imagem, uma vez que ambas parecem iguais. Muitos padrões da natureza, como rios, fendas, montanhas e nuvens, apresentam essa característica e são comumente invariantes à escala de observação (BUNDE; HAVLIN, 1994). Esses conceitos são explorados para análises de texturas em imagens. Isso ocorre porque alguns tipos de imagens, como a categoria de imagens médicas, não são apenas espectral e espacialmente complexas, mas frequentemente exibem certas semelhanças em diferentes escalas espaciais. A partir disso, a geometria fractal oferece a capacidade de descrever e caracterizar a complexidade das imagens ou, mais precisamente, de sua composição de textura (LOPES; BETROUNI, 2009).

2.2.1 Dimensão Fractal

A dimensão fractal (DF) é uma das principais medidas que descrevem a geometria fractal, caracterizando a complexidade e irregularidade de um fractal (IVANOVICI; RICHARD; DECEAN, 2009). Quando aplicada a uma superfície, se assemelha a noção intuitiva de rugosidade (PENTLAND, 1984). A DF pode ser vista como uma generalização, em termos de valores racionais e irracionais, da dimensão euclidiana, cujos valores se restringem somente aos números inteiros, como 1, 2, 3 ou n . Além disso, é uma grandeza livre de escala, capaz de quantificar o preenchimento fracionário de uma estrutura em algum intervalo de escala, mas que geralmente não se estende para escalas microscópicas ou macroscópicas infinitas (MANDELBROT, 1983). Existem diferentes maneiras de calcular o valor de dimensão fractal (HARRIS; STOCKER, 1998; HIGUCHI, 1988; IVANOVICI; RICHARD, 2011; NIKOLAIDIS; NIKOLAIDIS; TSOUROS, 2011), dependendo do tipo de imagem sob análise (CăLIMAN A.; IVANOVICI, 2012; NAYAK; MISHRA, 2016; RIBEIRO et al., 2019).

2.2.2 Lacunaridade

Quando a geometria fractal foi introduzida por Mandelbrot (MANDELBROT, 1977), existia a consciência de que a dimensão fractal não seria suficiente para capturar a complexidade dos fractais, uma vez que objetos com a mesma dimensão poderiam aparentar diferenças entre si. A lacunaridade (*Lac*) foi definida como uma métrica complementar, caracterizada por uma função de distribuição de massa. Isso permite indicar como o espaço está sendo ocupado, relacionando-se à organização topológica dos objetos em uma imagem (IVANOVICI; RICHARD; DECEAN, 2009).

Mais especificamente, a lacunaridade quantifica o desvio da invariância translacional de um objeto geométrico, permitindo indicar o quão similar entre si são as partes de diferentes regiões dos objetos. Dessa forma, imagens com baixa lacunaridade são mais homogêneas em relação à distribuição de tamanho e arranjo espacial das lacunas.

Também são invariantes por translação, visto que o tamanho das lacunas são iguais. Por outro lado, ao ter objetos com múltiplos tamanhos de lacunas, caracteriza-se uma heterogeneidade e a não invariância em termos de translação, representando uma alta lacunaridade (PLOTNICK; GARDNER; O'NEIL, 1993).

2.2.3 Percolação

A percolação foi definida por Broadbent e Hammersley (BROADBENT; HAMMERSLEY, 1957) e se refere, de maneira tradicional, ao movimento e filtragem de fluidos através de materiais porosos. Uma clássica representação é a passagem da água no pó de café (FONTES, 1996). A teoria da percolação é útil na caracterização de muitos dos sistemas desordenados, visto que o processo de percolação é puramente aleatório. Dessa forma, os locais do sistema em análise a serem ocupados ou que permanecerão vazios, são escolhidos de forma aleatória com probabilidade p , possibilitando a formação de aglomerados (STRELNICKER; HAVLIN; BUNDE, 2009). Quando um aglomerado forma um caminho que conecta as duas extremidades do sistema, então é considerado que houve percolação. A topologia obtida desse processo possui estruturas altamente relacionadas aos fractais. Esses conceitos também são aplicados na análise de imagens, considerando a percolação por meio da conectividade entre *pixels* vizinhos (YAMAGUCHI; HASHIMOTO, 2010).

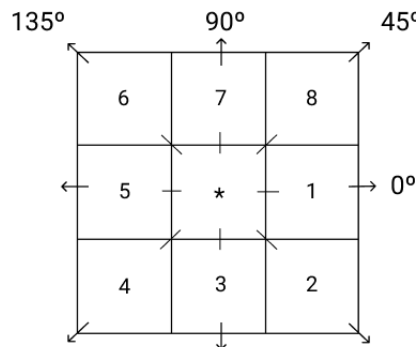
2.3 Descritores de Haralick

Os descritores de Haralick (HARALICK; SHANMUGAM; DINSTEN, 1973) conseguem quantificar informações sobre os padrões formados em texturas, sendo amplamente utilizados para classificar imagens (MIYAMOTO; MERRYMAN, 2005). Mais especificamente, estes descritores são estatísticas calculadas por meio de uma matriz de co-ocorrência. Esse tipo de matriz é frequentemente utilizada na área de

análise de imagens, pois representa características importantes da textura em uma dada região (TESAR et al., 2008).

Dada uma relação espacial definida entre *pixels* em uma textura, uma matriz de co-ocorrência P_{co} representa a distribuição de pares de *pixels* vizinhos em nível de cinza. O número de linhas e colunas da matriz é definida exclusivamente pela quantidade de níveis de cinza (N_g) da imagem (SIQUEIRA; SCHWARTZ W. R. ANS PEDRINI, 2013). A matriz é simétrica e possui dimensionalidade $N_g \times N_g$. Um elemento $P_{co}(m,n)$ da matriz indica o número de transições entre os níveis de cinza m e n que ocorrem na textura. Além disso, é possível considerar uma medida de distância δ e um ângulo arbitrário θ entre os pares em análise. Na Figura 2.1 estão representados ângulos específicos propostos por Haralick (HARALICK; SHANMUGAM; DINSTEIN, 1973) para o cálculo da matriz de co-ocorrência. Na Figura 2.2 é exemplificado o processo de construção de uma matriz de co-ocorrência com $\delta = 1$ e $\theta = 0^\circ$.

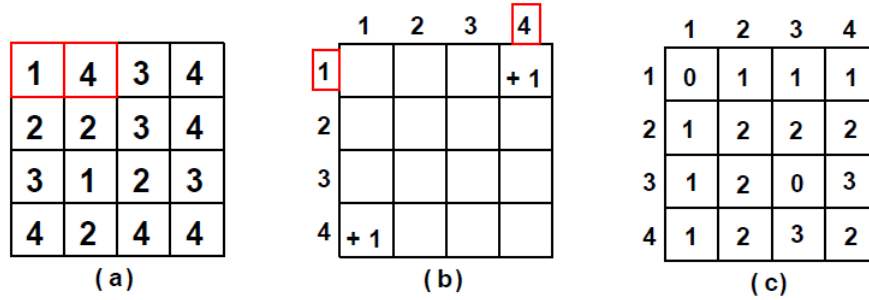
Figura 2.1: Ângulos θ considerados para o cálculo da matriz de co-ocorrência.



Fonte: (HARALICK; SHANMUGAM; DINSTEIN, 1973).

Uma vez calculadas as frequências para cada transição de escala de cinza, é possível extrair os descritores após a normalização da matriz. A normalização consiste em dividir o valor de cada elemento da matriz, $P_{co}(m,n)$, pela soma de todos os seus componentes, como definido na Equação 2.1:

Figura 2.2: Exemplificação do processo de construção da matriz de co-ocorrência, com $\delta = 1$ e $\theta = 0^\circ$: (a) Imagem original, com a indicação de pares de *pixels* vizinhos; (b) Estágio de incrementação do valor na matriz após analisar os *pixels* vizinhos destacados em (a); (c) Matriz de co-ocorrência resultante da imagem (a).



Fonte: (MIYAMOTO; MERRYMAN, 2005).

$$p_{m,n} = \frac{P_{co}(m,n)}{\sum_{i=0}^{H_g} \sum_{j=0}^{H_g} P_{co}(i,j)}, m, n = 0, \dots, H_g, \quad (2.1)$$

em que H_g define a maior intensidade de nível de cinza, m e n indicam os níveis de cinza e $p_{m,n}$ refere-se ao valor normalizado.

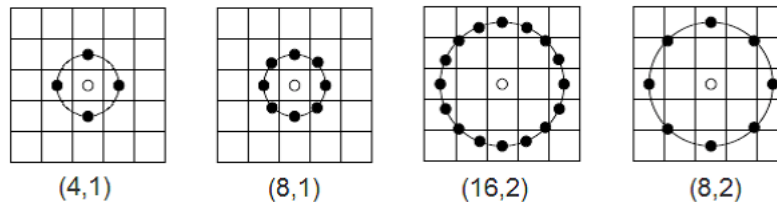
A partir da formulação apresentada previamente, é possível extrair 14 descritores propostos por Haralick (HARALICK; SHANMUGAM; DINSTEN, 1973): (1) Segundo momento angular; (2) Contraste; (3) Correlação; (4) Soma dos quadrados; (5) Momento inverso das diferenças; (6) Somatório da média; (7) Somatório da variância; (8) Somatório da entropia; (9) Entropia; (10) Diferença da variância; (11) Diferença da entropia; (12) e (13) são ambas medidas de informação de correlação; (14) Coeficiente de correlação maximal.

2.4 Local Binary Patterns – LBP

LBP é um descritor para texturas definidas em níveis de cinza a fim de computar a distribuição de padrões binários em uma vizinhança circular, tendo como referência um *pixel* central. O resultado é um código LBP para cada *pixel* da imagem (KORKMAZ; BINOL, 2018; OJALA; PIETIKÄINEN; HARWOOD, 1996). A vizinhança é

caracterizada por dois parâmetros: um raio R e um número de vizinhos PN (YU et al., 2019), como indicado na Figura 2.3.

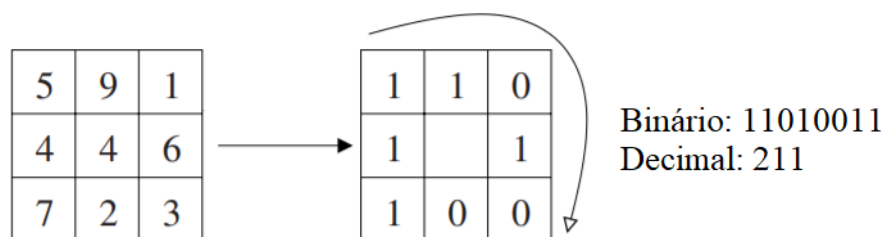
Figura 2.3: Diferentes configurações de vizinhança (PN, R) para o método LBP.



Fonte: (NGUYEN; VU; CAPLIER, 2012).

Se a intensidade do *pixel* vizinho for maior ou igual à intensidade de referência, é atribuído o valor binário 1. Caso contrário, atribui-se o valor 0, como indicado na Figura 2.4. Concatenando todos os valores da região em análise, a partir do canto superior esquerdo e em sentido horário, obtém-se um número binário. O valor decimal correspondente, nomeado por LBP ou código LBP, assume valores entre 0 a 255, sendo utilizado para representar a região analisada. O histograma das rotulagens LBP pode ser utilizado como um descritor de textura da imagem analisada (DORNAIKA et al., 2017).

Figura 2.4: Obtenção do código LBP, considerando $R = 1$ e $PN = 8$.



Fonte: Adaptado de (AHONEN; HADID; PIETIKÄINEN, 2004).

Há outra extensão do operador LBP, denominado de padrão uniforme. Um código LBP é denominado uniforme quando possuem no máximo duas transições 0-1 ou 1-0. Por exemplo, 00011110 e 11011111 são LBPs uniformes, enquanto 01010000 e 01001110 não o são. Seleccionando apenas os padrões uniformes, contribui-se para

a redução do comprimento do vetor de características, reduzindo os possíveis 256 valores para apenas 59, e melhorando o desempenho dos classificadores. Além disso, são invariantes à rotação (LAHDENOJA; POIKONEN J. ANTD LAIHO, 2013). O surgimento dessa variante foi motivada pelo fato de que padrões uniformes aparecem com mais frequência nas imagens do que os outros, correspondendo a pouco menos de 90% de todos os padrões quando utilizado uma vizinhança de $R = 1$ e $PN = 8$ (AHONEN; HADID; PIETIKÄINEN, 2004).

2.5 *Deep Learning*

Deep Learning, ou aprendizagem profunda, é uma subárea da inteligência artificial focada em treinar computadores para a realização de tarefas humanas. Nesta área, um conceito importante é o de aprendizado por representação ou por características: métodos que usam dados brutos, como os *pixels* de uma imagem, para descobrir automaticamente padrões para a detecção ou classificação de elementos. Os algoritmos de *deep learning* são compostos por múltiplos níveis de representação, obtidos pela composição de módulos simples não-lineares, que transformam as representações de um nível para uma representação ao nível superior, um pouco mais abstrata, começando pelos dados brutos de entrada (LECUN; BENGIO; HINTON, 2015). Essa hierarquia, composta por um número suficiente de camadas simples de transformações, permite que o computador aprenda conceitos complexos.

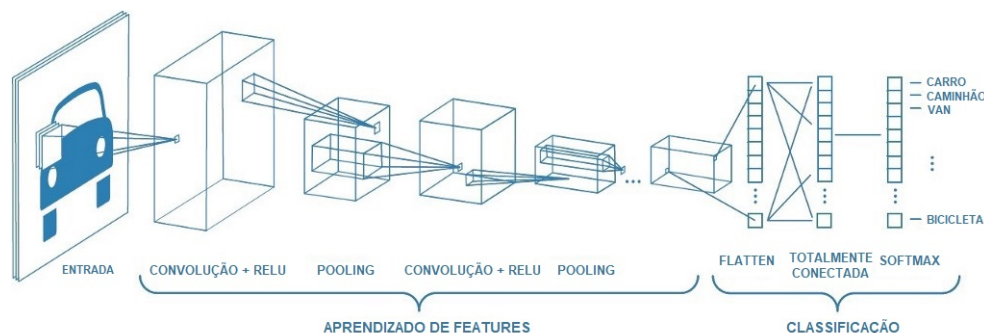
No contexto de análises de imagens, as características aprendidas pelos algoritmos *deep learning* são refinadas em camadas distintas. Por exemplo, as características aprendidas na primeira camada de representação, usualmente indicam a presença ou ausência de bordas em locais e orientações específicas na imagem. Na segunda camada, detectam-se os padrões por meio do reconhecimento de arranjos particulares das arestas, independentemente de pequenas variações nas posições das arestas. A terceira camada associa esses padrões em combinações maiores, correspondendo a

partes de objetos familiares, enquanto permite que as camadas subsequentes detectem objetos como combinações dessas partes (HAQUE; NEUBERT, 2020).

2.6 Redes Neurais Convolucionais

Um tipo particular de arquitetura *deep learning* é a Rede Neural Convolucional, do inglês *Convolutional Neural Network* (CNN) . Essa arquitetura foi projetada para processar dados no formato de múltiplos vetores ou matrizes como, por exemplo, uma imagem colorida composta por três matrizes que mapeiam as intensidades luminosas nos três canais de cores (LECUN; BENGIO; HINTON, 2015). Dessa forma, as camadas de uma CNN têm neurônios dispostos em três dimensões: largura, altura e profundidade. Assim como na maioria das redes neurais, uma CNN é constituída por neurônios interconectados que possuem entradas com pesos, *bias* e funções de ativação. Adicionalmente, uma CNN é composta por pilhas de diferentes camadas especializadas interconectadas (convolucional, ativação, *pooling*, totalmente conectada, *softmax* e outras), como ilustrado na Figura 2.5, cujos pesos são treinados pelo algoritmo de *backpropagation* (LECUN; BENGIO; HINTON, 2015). Nesse algoritmo, os pesos das camadas são atualizados, a partir da última camada até atingir a primeira delas, realizando uma retropropagação do erro obtido pela rede. As camadas mais profundas da CNN funcionam como extratores de características de baixo nível. As características obtidas a partir da última camada de uma rede pré-treinada estão sujeitas, no entanto, a uma forte dependência do conjunto de dados utilizados para treiná-la e do problema de classificação específico representado por um determinado conjunto de dados (NANNI; GHIDONI; BRAHNAM, 2018).

Figura 2.5: Exemplo de arquitetura CNN, considerando um problema de classificação de veículos.



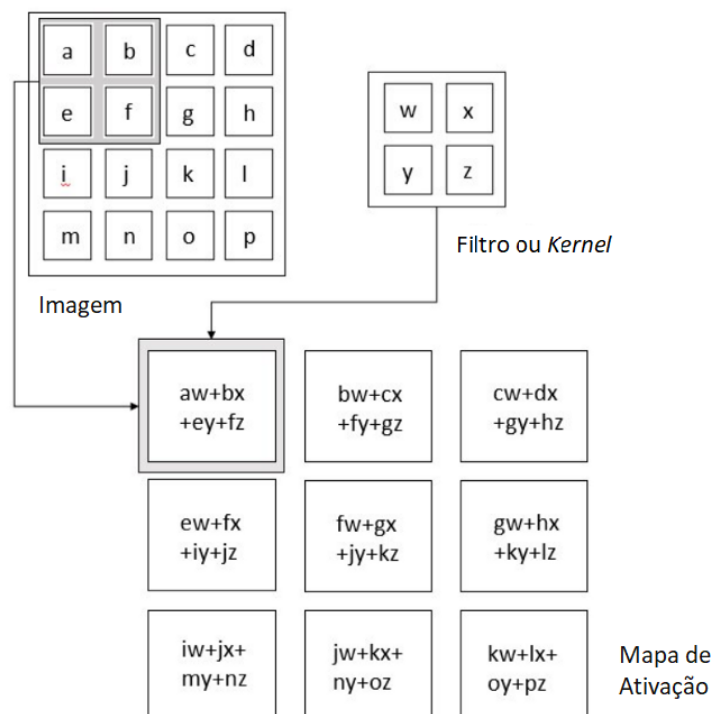
Fonte: Adaptado de (KULKARNI et al., 2019).

2.6.1 Camada Convolutiva

A arquitetura de uma rede neural convolucional considera uma organização típica de camadas convolucionais intercaladas com camadas de *pooling*, seguida pelas camadas totalmente conectadas (SHEN; WU; SUK, 2017). A camada convolucional tem seu nome baseado na operação que ela realiza: a convolução. O processo de convolução consiste em realizar o produto escalar entre duas matrizes: uma matriz é o conjunto de parâmetros aprendíveis, também denominados por filtros ou *kernel*; a outra matriz é o campo receptivo – uma parte restrita da imagem que representa um intervalo discreto, selecionado como entrada para o processo de convolução. O *kernel* é espacialmente menor que a imagem e a cada operação de convolução, desliza por toda a largura e altura da imagem. Em cada uma dessas operações, computa-se a soma do produto escalar entre o filtro e o campo receptivo, valor este que será a representação da imagem naquela região receptiva. Um exemplo do processo de convolução está ilustrado na Figura 2.6. Nela, o campo receptivo está sendo representado em cinza na imagem de tamanho 4×4 . O resultado da convolução entre o campo receptivo e o *kernel* produz um resultado no mapa de ativação (ou mapa de características), localizado na mesma região em que o campo receptivo se posiciona na imagem original.

As redes neurais convolucionais consistem em um conjunto de camadas que, por

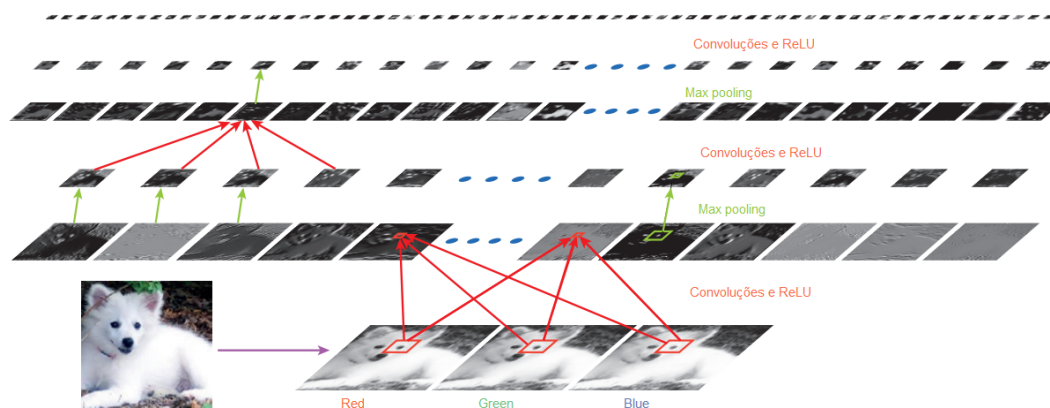
Figura 2.6: Exemplo do processo de convolução de uma imagem 4×4 com um filtro de tamanho 2×2 e o resultado da operação.



Fonte: Adaptado de (GOODFELLOW; BENGIO; COURVILLE, 2016).

sua vez, contém um ou mais planos. Cada unidade em um plano recebe informações de uma pequena vizinhança dos planos da camada anterior. Os pesos que formam os campos receptivos de um plano são forçados a serem iguais em todos os seus pontos. Cada plano, que é resultado de um processo de convolução na camada anterior, pode ser considerado um mapa de características que possui um detector de características fixo (Figura 2.7). Dessa forma, são normalmente utilizados múltiplos planos em cada camada, de forma que múltiplas características possam ser detectadas (LAWRENCE et al., 1997). Essas camadas que computam as saídas dos neurônios, que estão conectados às regiões locais, aplicando uma operação de convolução à entrada, são denominadas de camadas convolucionais e, apresenta-se como o bloco central na caracterização de uma CNN. As camadas convolucionais são responsáveis por tornar uma CNN computacionalmente custosa.

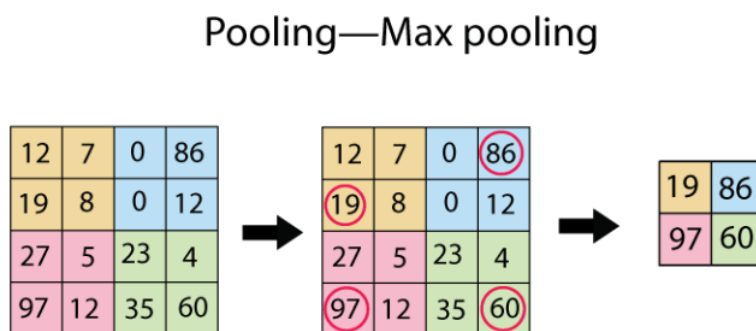
Figura 2.7: Saídas de cada camada (horizontal) de uma CNN. Cada imagem é um mapa de características correspondente à saída das características aprendidas na camada anterior, detectada em cada posição da imagem.



Fonte: Adaptado de (LECUN; BENGIO; HINTON, 2015).

2.6.2 Pooling

A camada de *pooling*, comumente posterior à camada de convolução, realiza operações não-lineares de subamostragem, que reduz a resolução dos mapas de características obtidos pela camada convolucional anterior. Mais especificamente, cada mapa de características da camada de *pooling* é vinculada a um mapa de características da camada convolucional. Cada valor do mapa de características da camada de *pooling* é calculada com base em um subconjunto de valores dentro do campo receptivo local do mapa de características convolucional correspondente. Semelhante à camada convolucional, o campo receptivo encontra um valor representativo entre as unidades em seu campo (SHEN; WU; SUK, 2017). A estratégia mais utilizada para determinar esse valor é operação *max pooling*. Nela, a entrada é particionada em um conjunto de retângulos não sobrepostos e o valor significativo produzido em cada um deles será aquele que apresentar o maior valor. Um exemplo da operação de *max pooling* é ilustrado na Figura 2.8. Dessa forma, enquanto há redução espacial, também há a redução do número de parâmetros, da possibilidade de *overfitting* e da complexidade da rede (IDE; KURITA, 2017). O papel desempenhado pela camada de *pooling* é mesclar as características semelhantes em apenas uma.

Figura 2.8: Operação de *max pooling*.

Fonte: (WILL, 2018).

2.6.3 Função de Ativação

Antes da camada de *pooling* receber a saída da camada convolucional, os dados são aplicados a uma função de ativação. As funções de ativação são utilizadas para computar a soma ponderada das entradas e *bias*, decidindo se um neurônio da rede será ativado ou não. As funções de ativação são elementos cruciais na arquitetura CNN, pois transformam um classificador linear em um não-linear, o que provou ser um ponto-chave para o alto desempenho em uma ampla gama de tarefas nos últimos anos (EGER; YOUSSEF; GUREVYCH, 2018). Uma das funções de ativação mais utilizadas atualmente é a ReLU (Unidade Linear Retificada), que calcula o maior valor entre 0 e o valor dado como entrada:

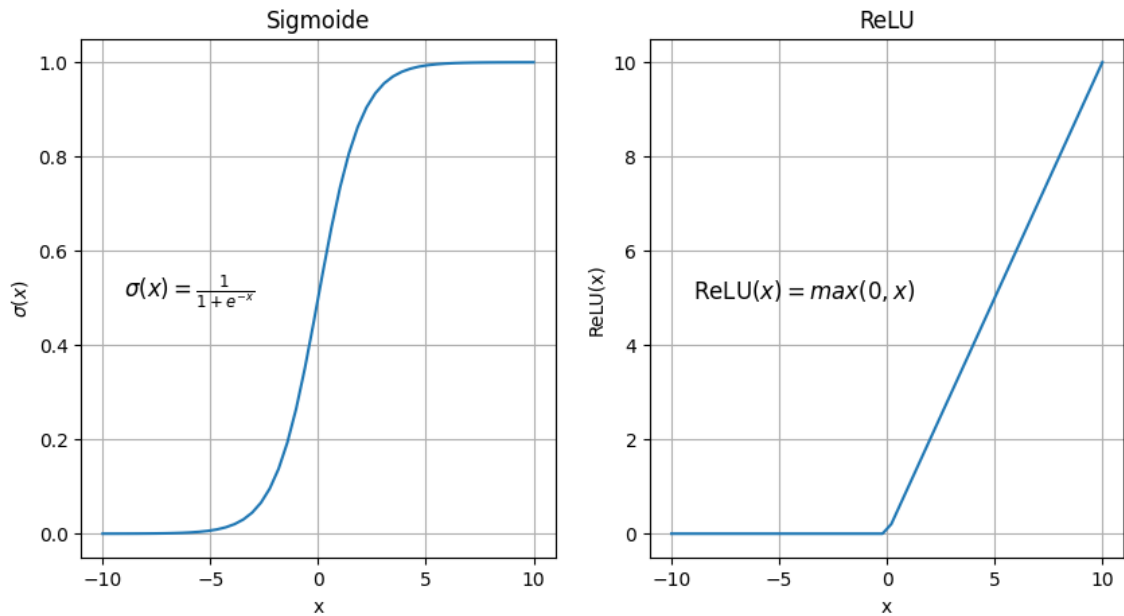
$$ReLU(x) = \max(0, x). \quad (2.2)$$

Essa função tem como principal vantagem sua rápida computação, dado que não há a necessidade de realizar operações como exponenciação ou divisão.

Além da ReLU, outra função bem conhecida é a Sigmoide, aplicada com sucesso em problemas de classificação binária. Os valores assumidos pela função variam entre 0 e 1, conforme dado na Equação 2.3. Na Figura 2.9 é ilustrado o comportamento dessas duas funções, ReLU e Sigmoide.

$$\text{Sigmoid}(x) = \frac{1}{1 + \exp(-x)}. \quad (2.3)$$

Figura 2.9: Ilustrações de funções de ativação: Sigmoide, à esquerda; ReLU, à direita.



Fonte: Elaborado pelo autor.

Por fim, a função de ativação *SoftMax* é utilizada na abordagem de modelos multiclases, retornando uma probabilidade para cada classe do problema. Essa função é comumente aplicada nas camadas de saída das arquiteturas de aprendizado profundo, retornando como saída valores entre 0 e 1, cuja soma das probabilidades de todas as classes é igual a 1. A classe buscada é aquela com a maior probabilidade. As probabilidades são calculadas pela seguinte função (NWANKPA et al., 2018):

$$\text{SoftMax}(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)}. \quad (2.4)$$

2.6.4 Transferência de Aprendizado

Um dos principais problemas enfrentados pelas abordagens *deep learning* consiste na quantidade de dados necessários para treinar uma rede. Uma solução para esse

problema é a aplicação do aprendizado por transferência, ou *transfer learning*. Essa técnica baseia-se em transferir o conhecimento aprendido em uma ou mais tarefas de origem e usá-lo para melhorar a aprendizagem em uma tarefa de destino relacionada (TORREY; SHAVLIK, 2009).

As arquiteturas CNN pertencem à classe de algoritmos de aprendizagem indutiva, cujo objetivo consiste em mapear as características de entrada entre as classes para obter uma generalização dos dados. Assim, a aprendizagem indutiva pode ser transferida de uma arquitetura treinada na tarefa fonte para a classe alvo, realizada por meio do ajuste do espaço do modelo e corrigindo os valores de *bias*. Usualmente, isso pode ser realizado substituindo a última camada do modelo (classificador), que deve ser treinado nas características generalizadas (ALMARAZ-DAMIAN et al., 2020), proporcionando menor tempo de treinamento e suprimindo a possibilidade de *overfitting*, caso a quantidade de dados não fosse suficiente.

2.7 Arquiteturas de Redes Neurais Convolucionais

Diferentes arquiteturas foram propostas ao longo dos anos, desde modelos menos profundos, com oito camadas convolucionais, como a arquitetura AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), como também modelos mais profundos, como a EfficientNet (TAN; LE, 2019), com mais de 237 camadas treináveis. Os modelos de redes neurais apresentados nas próximas subseções apresentaram resultados relevantes quando considerado o problema de classificação de imagens médicas (NANNI et al., 2020; ROBERTO et al., 2021; TRIPATHI; SINGH, 2020), além de fornecerem as melhores colocações na competição “*ImageNet Large Scale Visual Recognition Challenge*” (ILSVRC), cujo intuito foi a avaliação dos algoritmos para detecção de objetos e classificação de imagens no banco de imagens ImageNet (RUSSAKOVSKY et al., 2015).

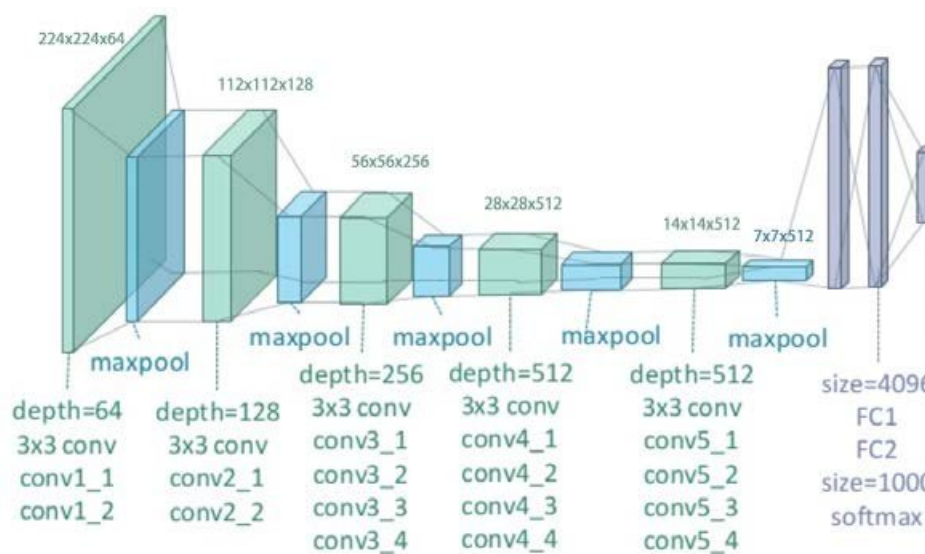
2.7.1 VGG-19

A arquitetura VGGNet foi proposta em 2014 (SIMONYAN; ZISSERMAN, 2014), obtendo a segunda posição no ILSVRC 2014, perdendo somente para a arquitetura GoogleNet (SZEGEDY et al., 2015a). Dois modelos foram propostos, um com 16 camadas treináveis e outro com 19 camadas treináveis, justificando os nomes VGG-16 e VGG-19, respectivamente. A arquitetura surgiu com a ideia de utilizar filtros menores em redes mais profundas, de no mínimo 12 camadas convolucionais, resultando em cerca de 143 milhões de parâmetros para o modelo VGG-19. A arquitetura é dividida em cinco blocos compostos por camadas convolucionais conectadas, sem qualquer outra camada entre elas, seguida por uma camada de *max pooling*. Os primeiros dois blocos possuem duas camadas convolucionais, enquanto o restante é constituído por quatro camadas. Ao final, três camadas totalmente conectadas são utilizadas na classificação, com as duas primeiras compostas por 4096 neurônios e a última com o total de neurônios relativo à quantidade de classes analisadas. Em toda a rede são considerados *kernels* de tamanhos 3×3 e *stride* de 1 *pixel*. A camada de *max pooling* realiza as operações com uma janela de tamanho 2×2 com *stride* de 2 *pixels*. Em relação às funções de ativação das camadas ocultas, utiliza-se a função ReLU, enquanto na camada de saída, a função *SoftMax* (SIMONYAN; ZISSERMAN, 2014). Uma ilustração da rede é apresentada na Figura 2.10.

2.7.2 ResNet-50

A arquitetura ResNet-50, vencedora da competição ILSVRC 2015, é uma rede baseada em blocos residuais (HE et al., 2016). Um exemplo de bloco residual é ilustrado na Figura 2.11. Na imagem, a saída de x passa pela conexão *shortcut path* (ou *skip connections*), a fim de que seja somada com o caminho principal $f(x)$. É importante destacar que $f(x)$ consiste na saída da última camada convolucional do bloco, exatamente antes da função de ativação ReLU. Logo, a saída do bloco residual será

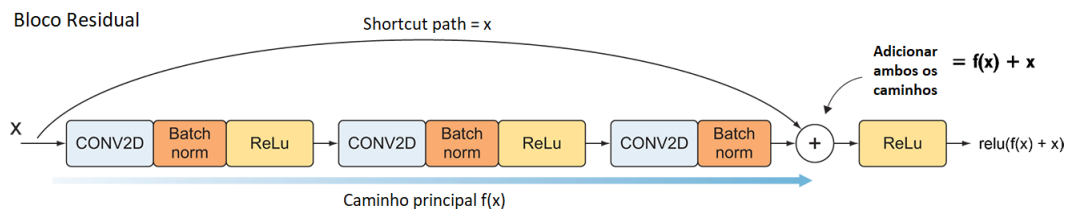
Figura 2.10: Arquitetura do modelo VGG-19.



Fonte: (ZHENG; YANG; MERKULOV, 2018).

$$\text{ReLU}(f(x) + x).$$

Figura 2.11: Ilustração de um bloco residual na arquitetura ResNet-50.



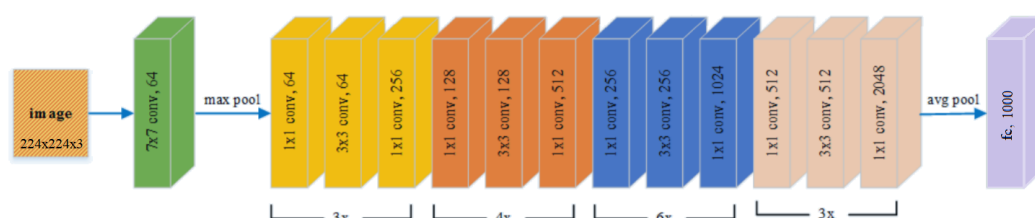
Fonte: Adaptado de (ELGENDY, 2020).

A ideia que fundamenta a arquitetura ResNet é utilizar as conexões residuais para solucionar o problema de degradação ao treinar uma rede muito profunda, permitindo que o gradiente seja retropropagado diretamente para as camadas anteriores. Outro benefício importante das *skip connections* é permitir que o modelo aprenda uma função de identidade para garantir que a camada terá um desempenho pelo menos tão bom quanto a camada anterior (ELGENDY, 2020).

O modelo ResNet-50 é composto por cinco blocos convolucionais empilhados (MAHMOOD et al., 2020). Na Figura 2.12 está indicado um diagrama referente à rede. Cada bloco residual da arquitetura é repetido em uma quantidade de vezes in-

dicada inferiormente. As camadas de um mesmo bloco residual estão indicadas pela mesma cor na Figura 2.12. Assim, o primeiro bloco convolucional é composto por uma única camada convolucional com 64 *kernels* de dimensões 7×7 . O segundo bloco da rede (em amarelo) é composto por três blocos residuais. Os três blocos seguintes são formados por quatro, seis e três blocos residuais. Após o quinto bloco convolucional, anteriormente à camada totalmente conectada, é aplicada uma camada de *average pooling* (HE et al., 2016).

Figura 2.12: Arquitetura ResNet-50.



Fonte: Adaptado de (TALO, 2019).

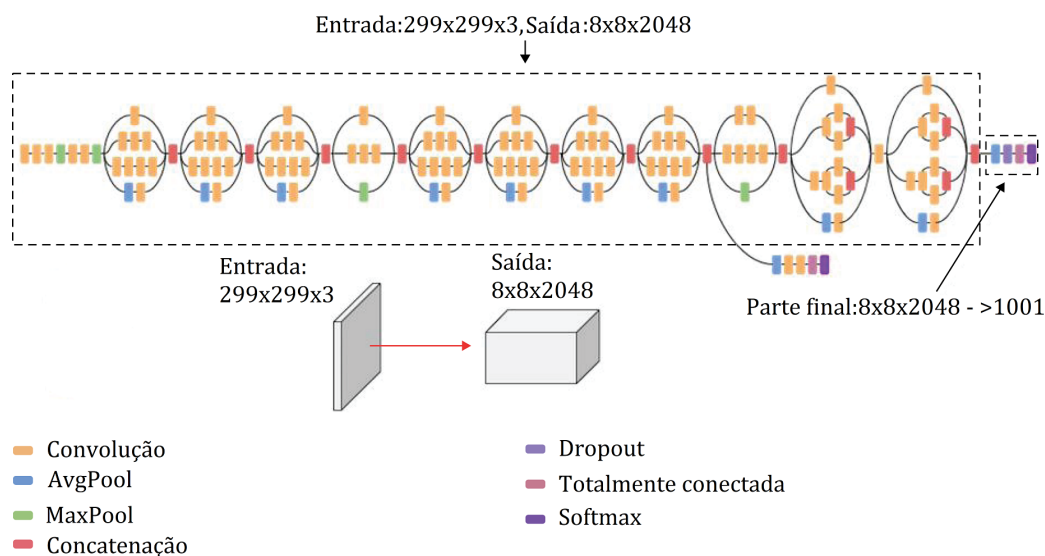
2.7.3 Inception v3

A arquitetura Inception v3 foi proposta em 2015 (SZEGEDY et al., 2015b), contando com 48 camadas de profundidade, como ilustrado na Figura 2.13. Essa rede é a terceira da família Inception, possuindo algumas melhorias em relação às anteriores. Ela foi projetada para funcionar bem mesmo sob restrições de memória ou de custo computacional, superando os desempenhos obtidos via arquitetura VGGNet (HE et al., 2015).

Alguns pontos importantes que constituem esse modelo são (SZEGEDY et al., 2015b):

- **Fatorar Convoluções:** aqui foram exploradas formas de aumentar a eficiência computacional do modelo, isso porque qualquer redução no custo computacional resulta em uma redução do número de parâmetros e um treinamento

Figura 2.13: Arquitetura da rede Inception v3.



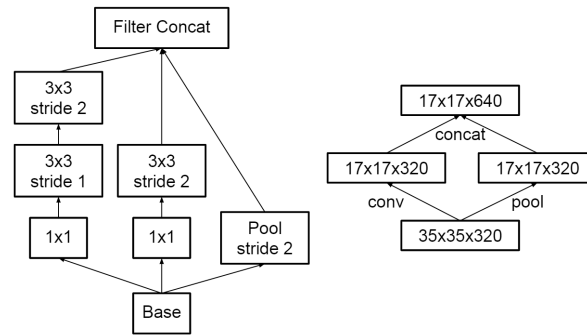
Fonte: Adaptado de (RAMANESWARAN et al., 2021).

mais rápido. Para isso, os autores indicaram que, por exemplo, uma convolução 5×5 tem um gasto computacional 2,78 vezes maior do que dois filtros de 3×3 . Portanto, em termos de fatoração, uma convolução é dividida em convoluções menores;

- **Convoluções Assimétricas:** considerando os pontos abordados anteriormente, os autores propuseram substituir as convoluções $n \times n$ por uma convolução $1 \times n$, seguida por outra convolução $n \times 1$. Dessa forma, a economia de custo computacional aumenta drasticamente à medida que n cresce;
- **Classificadores Auxiliares:** classificadores auxiliares foram colocados entre as camadas da rede durante a fase de treinamento. A função de perda foi adicionada à função principal da rede. Esses classificadores atuam como regularizadores, já que o classificador principal da rede funciona melhor se a ramificação lateral for normalizada em lote (*batch-normalization*) ou tiver uma camada *dropout*;
- **Redução do tamanho do *grid*:** tradicionalmente, para a redução do tamanho dos mapas de características (*grid*), foram utilizadas operações de *pooling*.

Nesse modelo, para reduzir o tamanho da rede eficientemente, a dimensão de ativação dos filtros de rede foi expandida. Isso quer dizer que, por exemplo, dado um *grid* $d \times d$ com k filtros, se quisermos chegar a um *grid* de dimensões $\frac{d}{2} \times \frac{d}{2}$ com $2k$ filtros, será aplicada uma convolução com $2k$ filtros com *stride* 1 e, um passo adicional com uma operação de *pooling*. Na Figura 2.14 é ilustrado esse processo.

Figura 2.14: Redução do tamanho do *grid*. O diagrama na direita representa a mesma solução à esquerda, mas da perspectiva de tamanho do *grid* ao invés de operações.



Fonte: (SZEGEDY et al., 2015b).

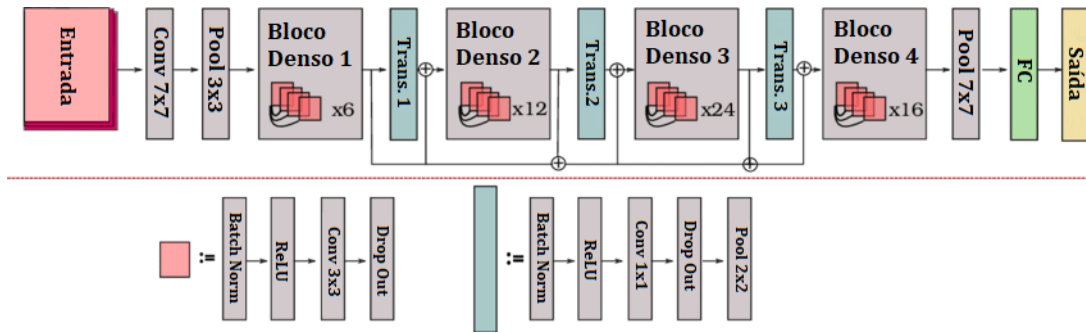
2.7.4 DenseNet-121

A arquitetura DenseNet-121 (HUANG et al., 2017) se diferencia por conectar cada camada a todas as outras camadas de forma *feed-forward*, formando os chamados *dense block* (HUANG et al., 2017). Assim, para cada camada do bloco, os mapas de características de todas as camadas anteriores são utilizadas como entrada e seus próprios mapas de características serão utilizados como entrada às camadas subsequentes.

A família de arquiteturas DenseNet é composta por modelos de diferentes profundidades, sendo a DenseNet-121 formada por 121 camadas treináveis: 117 camadas convolucionais; 3 camadas de transição; 1 camada de classificação. Cada camada convolucional corresponde, na verdade, a uma sequência de operações: *batch normalization*-ReLU-Camada convolucional (LI et al., 2020). Uma transição é caracterizada

por uma camada convolucional 1×1 e uma camada de *avgpool* 2×2 com *stride* 1. A etapa de classificação da rede contém uma camada global de *avgpool* 7×7 , seguida por uma camada totalmente conectada e uma *softmax*. A arquitetura possui quatro blocos densos. Na Figura 2.15 é ilustrado o modelo descrito. Entre as vantagens de utilizar esse modelo estão: minimizar o problema de dissipação do gradiente, fortalecer a propagação das *features*, incentivar a reutilização de *features* e reduzir substancialmente o número de parâmetros (HUANG et al., 2017).

Figura 2.15: Arquitetura DenseNet-121.



Fonte: Adaptado de (RADWAN, 2019).

2.7.5 EfficientNet

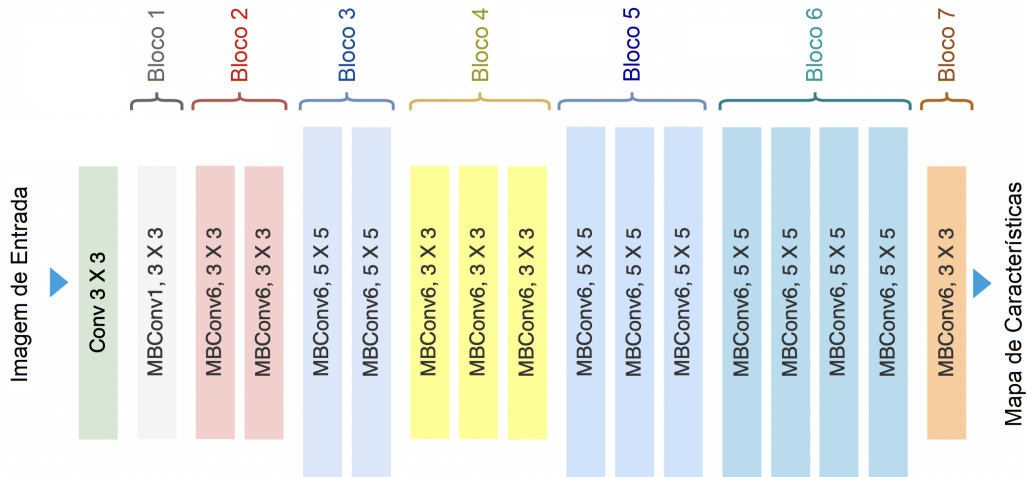
A família de arquiteturas EfficientNet surgiu com o intuito de escalar redes convolucionais nas três dimensões – largura, profundidade e resolução – de maneira uniforme e efetiva, tomando como base um conjunto fixo de coeficientes de escalonamento (TAN; LE, 2019). Isso significa que, se quisermos utilizar 2^N vezes mais recursos computacionais, então bastaria aumentar a profundidade da rede por α^N , a largura por β^N e a resolução da imagem por γ^N , considerando α, β e γ como os coeficientes constantes determinados por um processo de *grid search* na arquitetura base, a EfficientNet-B0.

Para escalar uniformemente todas as dimensões, os autores propuseram um modelo que utiliza um coeficiente composto ϕ , como descrito a seguir:

$$\begin{aligned}
\text{profundidade} &= \alpha^\phi \\
\text{largura} &= \beta^\phi \\
\text{resolução} &= \gamma^\phi \\
\text{s.t. } \alpha \cdot \beta^2 \cdot \gamma^2 &\approx 2 \\
\alpha \geq 1, \beta \geq 1, \gamma \geq 1
\end{aligned} \tag{2.5}$$

A partir da EfficientNet-B0 foram geradas as arquiteturas da família, desde a EfficientNet-B1 até a EfficientNet-B7. Para realizar esse processo foram encontrados os melhores parâmetros de α, β e γ na arquitetura B0, fixando $\phi = 1$. Após encontrar esses valores, os autores fixaram as constantes encontradas e escalaram o modelo variando o valor de ϕ , segunda a Equação 2.5. Uma representação da arquitetura EfficientNet-B0 é mostrada na Figura 2.16.

Figura 2.16: Arquitetura EfficientNet-B0.



Fonte: Adaptado de (AHMED; SABAB, 2020).

A base para o modelo EfficientNet é a camada MBConv, um bloco residual com estrutura invertida, aplicado na rede MobileNetV2. Cada arquitetura possui semelhanças com suas versões anteriores. O que os diferenciam são os diferentes mapas de carac-

terísticas que aumentam o número de parâmetros. Todos os modelos têm a mesma arquitetura que a sua versão anterior, porém, com variações no número de sub-blocos, que se expandem conforme se avança da EfficientNet-B0 para a EfficientNet-B7 (AHMED; SABAB, 2020).

2.8 Técnicas para Seleção de Características

O aumento de características não implica, necessariamente, em obter maior discriminação dos grupos sob investigação. Vetores de características com alta dimensionalidade retardam o processo de aprendizado, causando também o comprometimento do modelo de generalização e a tendência de *overfitting* do classificador (KOTSIANTIS, 2011). Para contornar essa problemática, a seleção de características se mostra uma estratégia amplamente eficaz. Com o foco em selecionar um subconjunto de atributos a partir do conjunto de entrada, a seleção de características possibilita reduzir os efeitos de ruídos ou atributos irrelevantes, reduzir o tempo de treinamento do modelo e melhorar o desempenho da predição, da visualização e do entendimento dos dados (CHANDRASHEKAR; SAHIN, 2014).

Um dos grupos de algoritmos para seleção de características é o *Wrapper*. Esse grupo é caracterizado pela utilização de algoritmos de busca e de um preditor, que opera como função objetivo para avaliar os desempenhos de subconjuntos selecionados. Inúmeros algoritmos de busca podem ser utilizados para encontrar um subconjunto de atributos que maximiza a função objetivo. Porém, como o número de subconjuntos a serem avaliados podem possuir tamanho considerável, uma busca exaustiva demandaria amplo recurso computacional. Dessa forma, pode-se recorrer ao uso das meta-heurísticas, pois produzem bons resultados e são computacionalmente viáveis.

As meta-heurísticas constituem algoritmos projetados para resolverem problemas de otimização sem a necessidade de se adaptar profundamente ao problema em questão. Geralmente, estas técnicas são aplicadas quando não há um algoritmo específico e

satisfatório para resolver um problema. As meta-heurísticas, por exemplo, podem ser pautadas em princípios físicos, biológicos ou etológicos (BOUSSAÏD; LEPAGNOT; SIARRY, 2013). Um algoritmo desta categoria tem êxito quando consegue balancear exploração e exploração. É necessária a exploração para identificar possíveis soluções em todo o espaço de busca. A exploração é importante para intensificar as buscas pelas melhores soluções encontradas até o momento. Algumas meta-heurísticas são baseados em:

- **Algoritmo Genético.** De forma análoga à genética de populações, os algoritmos genéticos realizam as buscas em meio a várias gerações de uma população. O critério de sobrevivência do mais apto fornece uma pressão evolutiva para que as populações desenvolvam indivíduos cada vez mais aptos (LIEPINS; VOSE, 1990). A população é composta por possíveis soluções do problema (indivíduos), codificadas mediante cromossomos artificiais. Dessa forma, esses cromossomos são submetidos a processos evolucionários, envolvendo avaliação, seleção, *crossover* e mutação (PACHECO, 1999). O ciclo básico de um algoritmo genético pode ser generalizado como:
 1. Avaliar a aptidão do indivíduo;
 2. Formar a população escolhendo indivíduos baseados na sua aptidão;
 3. Recombinação e mutação dos indivíduos.
- **Enxame de Partículas (*Particle Swarm Optimization*).** O algoritmo de enxame de partículas, PSO, proposto por (KENNEDY; EBERHART, 1995) simula o comportamento social de animais, como os bandos de pássaros ou cardumes, baseado em princípios sociopsicológicos. É um algoritmo apropriado para problemas de seleção de características, pois apresenta características que facilitam a representação dos atributos e busca global, além de utilizar poucos parâmetros (IFTIKHAR, 2015). No algoritmo PSO, uma população, denominada de

enxame, é composta por várias partículas. Cada uma dessas partículas são possíveis soluções que vagam no espaço de busca definido pelo problema, tendo associada a si valores de posição e velocidade inicializados de forma aleatória. A cada iteração, as partículas se movem na busca de uma solução ótima. Para isso, é realizada a atualização da sua posição baseando-se tanto na sua experiência, quanto na das partículas vizinhas (XUE; ZHANG; BROWNE, 2014). Como muitas partículas realizam esse tipo de comportamento simultaneamente, com uma tendência de agrupamento em regiões ideais do espaço, denomina-se a eles o conceito de enxame de partículas;

- **Comportamento Animal.** Algoritmos baseados em comportamento animal abordam o conceito de organização individual para entender o comportamento coletivo dos animais. Pressupõe-se como princípio central da organização individual que interações simples repetidas entre os indivíduos acarretam padrões comportamentais complexos ao nível de grupo. Essa inspiração se originou a partir dos padrões de comportamentos observados anteriormente em vários grupos de animais. Dessa forma, os algoritmos assumem a existência de um conjunto de operações que se assemelham às regras de interações que modelam o comportamento coletivo desse grupo de animais. Cada possível solução no espaço de busca é representada por um animal. O valor de aptidão pode ser considerado a dominância do animal no grupo, podendo ainda ser considerada a existência de uma memória coletiva. O processo completo do algoritmo imita o comportamento coletivo dos animais em questão (CUEVAS et al., 2012).

Um segundo grupo de seletores é o nomeado *filter*. Esses algoritmos (de filtragem) utilizam métodos de ranqueamento como principal critério para a seleção das características, resultando na ordenação dessas. Usualmente, um critério de ranqueamento adequado é aplicado para pontuar as características e um ponto de corte é definido para descartar aquelas abaixo desse limite. O uso desse seletor é justificado

principalmente devido à sua simplicidade e aos bons resultados em aplicações práticas (CHANDRASHEKAR; SAHIN, 2014). Um exemplo de algoritmo da categoria é o ReliefF.

A ideia desse algoritmo é estimar a qualidade dos atributos baseado no quão bem seus valores se distinguem entre instâncias que estão próximas umas às outras. Durante a etapa de treinamento do algoritmo, as instâncias são selecionadas aleatoriamente. Em seguida, o algoritmo busca outros dois conjuntos de valores: os k vizinhos mais próximos da mesma classe, nomeados por *nearest hits*; e os k vizinhos mais próximos de cada uma das diferentes classes, denominados *nearest misses*. Os valores finais são estimados baseados na distância de Manhattan e na média da contribuição de todos os *hits* e *misses* (LE et al., 2019; ROBNIK ŠIKONJA; KONONENKO, 2003).

2.9 Classificadores

O processo de classificação consiste na atribuição dos dados de entrada em uma ou mais classes pré-especificadas, com base nos processos de extração de características significativas e no processamento ou análise desses atributos. Os algoritmos que realizam esse mapeamento são denominados de classificadores ou algoritmos de classificação. Os classificadores são definidos como supervisionados quando mapeiam os dados de acordo com classes previamente rotuladas (PEDRINI; SCHWARTZ, 2008).

Para que as amostras sejam classificadas, é realizada uma etapa de treinamento anteriormente à fase de classificação. Portanto, é necessário um conjunto de treinamento formado por amostras representativas (e geralmente rotuladas) de tipos de padrões a serem encontrados na aplicação real. De maneira geral, o conjunto de treinamento provê informação significativa de como associar os dados de entrada com as decisões de saída, permitindo ao sistema aprender informações relevantes (WAH, 2008).

Os classificadores podem ser divididos em classes que caracterizam o seu funcionamento. Algumas dessas classes são:

- **Baseado em Função.** Em geral, esses tipos de classificadores utilizam estratégias como, por exemplo, a separação dos dados por meio de funções, como o classificador *Support Vector Machine* (SVM). Nele, o classificador aprende baseado na projeção do conjunto de treinamento em um espaço de características de alta dimensão, no qual duas classes estão separadas por um hiperplano que maximiza a margem de separação entre eles. Uma amostra não rotulada é classificada conforme o lado do plano onde ela é mapeada (KUMAR et al., 2020). A função que separa as classes é denominada de *kernel*. Outro classificador que se encaixa na categoria é a regressão logística. Baseado no método de regressão, é utilizado para prever uma variável dependente dicotômica (binária). SVM é um classificador útil para situações em que se deseja prever a presença ou a ausência de uma característica com base em valores de um conjunto de variáveis preditoras. É como um modelo de regressão linear, mas é adequado para modelos em que a variável dependente é dicotômica (KURT; TURE; KURUM, 2008). Existe, no entanto, o modelo de regressão logística multinomial, sendo uma generalização da regressão logística para problemas de várias classes (KARSMARKERS; PELCKMANS; SUYKENS, 2007);
- **Baseado em Árvore de Decisão.** A árvore de decisão particiona os dados hierarquicamente em uma estrutura de árvore, associando diferentes partições ao nível das folhas às diferentes classes. As partições são criadas de acordo com um critério de divisão: utilizar uma condição ou predicado em um único atributo, ou uma condição para vários atributos. De modo geral, o objetivo é dividir recursivamente os dados de treinamento de modo a maximizar a discriminação entre as diferentes classes em diferentes nós (AGGARWAL, 2015). Um exemplo de classificador baseado em árvore de decisão é o *Random Forest*. Nesse classificador é utilizado uma combinação de várias árvores de decisão. Cada árvore apresenta uma predição de classe. A classe mais predita será a classe

considerada na classificação final (BREIMAN, 2001);

- **Baseado em *Lazy Learning*.** Também conhecido por aprendizagem baseada em instâncias, os algoritmos baseados em *lazy learning* armazenam as instâncias de treinamento e esperam pelas instâncias de teste. Quando um dado de teste é recebido, a classificação ocorre considerando que essa amostra pertencerá à mesma classe da instância armazenada mais próxima, ou seja, a classe do dado de treinamento mais próximo àquela nova instância (AHA, 1997). O classificador KNN (*K-Nearest Neighbors*) é um exemplo dessa categoria. O algoritmo é baseado em uma função de distância que calcula a diferença ou similaridade entre duas amostras. Comumente é utilizada a distância Euclidiana entre duas instâncias. A partir disso, para uma dada amostra no espaço de busca, o algoritmo KNN encontra os K vizinhos mais próximos (*Nearest Neighbors*) e a classe mais comum entre eles. Essa classe será então atribuída à amostra sob análise (JIANG et al., 2007).
- **Baseados no Teorema de Bayes.** Classificadores como o *Naive Bayes* pertencem à família de classificadores probabilísticos baseados na aplicação do teorema de Bayes, com a suposição de que as características são fortemente independentes, indicando o nome de bayesiano ingênuo (*Naive Bayes*) (GRANIK; MESYURA, 2017). O classificador é treinado de forma que cada instância x é descrita como um conjunto de valores que pode assumir qualquer valor em um conjunto finito, mapeada por uma função $f(x)$. Um conjunto de exemplos de treinamento da função de destino é fornecido e uma nova instância é apresentada. O classificador bayesiano atribui à amostra a classe provável, segundo o cálculo da probabilidade condicional, baseando-se no conjunto de treinamento (KANTARCIOGLU; VAIDYA, 2003).

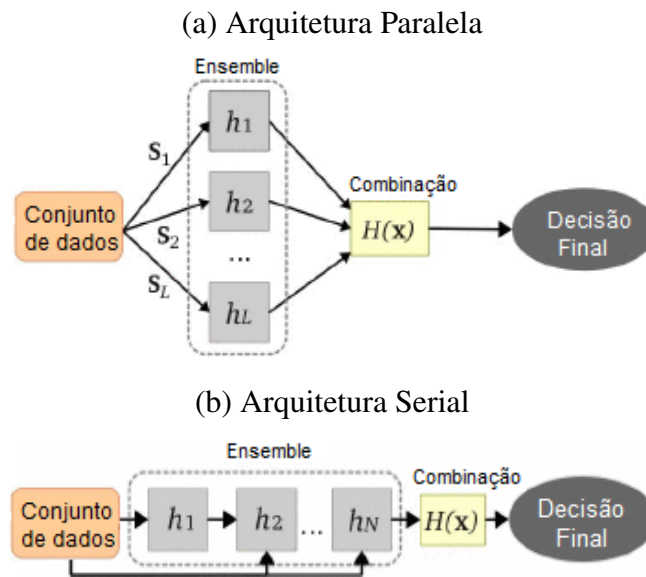
2.10 *Ensemble Learning*

O conceito de *ensemble learning* pode ser encontrado sob diversas terminologias: comitê, fusão de classificadores, múltiplos classificadores, combinação, agregação, mistura de especialistas ou outros termos que indicam um conjunto de algoritmos de aprendizado trabalhando cooperativamente. A ideia é as decisões serem combinadas para melhorar o desempenho do sistema de reconhecimento de padrões ou tornar as saídas mais confiáveis, por exemplo, sem *overfitting*. Muitos estudos se concentram em melhorar a precisão de problemas difíceis, gerenciando as fraquezas e pontos fortes de cada modelo, a fim de dar a melhor decisão possível considerando todo o conjunto (PONTI JR., 2011). A partir de estudos, verificou-se que problemas de classificação costumam ser mais precisos quando usado um *ensemble* de classificadores, ao invés de um classificador base individual (HAGERTY et al., 2019; ROBERTO et al., 2021). Assim, classificadores “fracos” podem fornecer melhores desempenhos do que aqueles altamente específicos (KUNCHEVA; BEZDEK; DUIN, 2001).

Neste contexto, considerando h_i como o i -ésimo classificador em um *ensemble* de L classificadores, $S = (x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$ o conjunto de treinamento, S_i uma versão do conjunto de dados disponível e, $H(x)$ uma função que combina/seleciona as decisões de vários algoritmos de aprendizagem sobre um novo padrão de entrada x , um *ensemble* pode ser construído de duas formas principais, serial ou paralelo. Na arquitetura paralela (Figura 2.17(a)), os classificadores são treinados em paralelo, de modo que suas saídas são combinadas posteriormente para obter a decisão final. Na arquitetura serial (2.17(b)), um classificador primário é usado e, quando não é possível classificar algum novo padrão, rejeitando-o, utiliza-se um segundo classificador treinado para ser preciso quanto aos erros do classificador anterior. Um terceiro e um quarto classificador podem ser usados e, assim por diante (PONTI JR., 2011).

Segundo (KUNCHEVA, 2001), existem quatro principais abordagens que podem ser adotadas a fim de criar o *ensemble* (Figura 2.18):

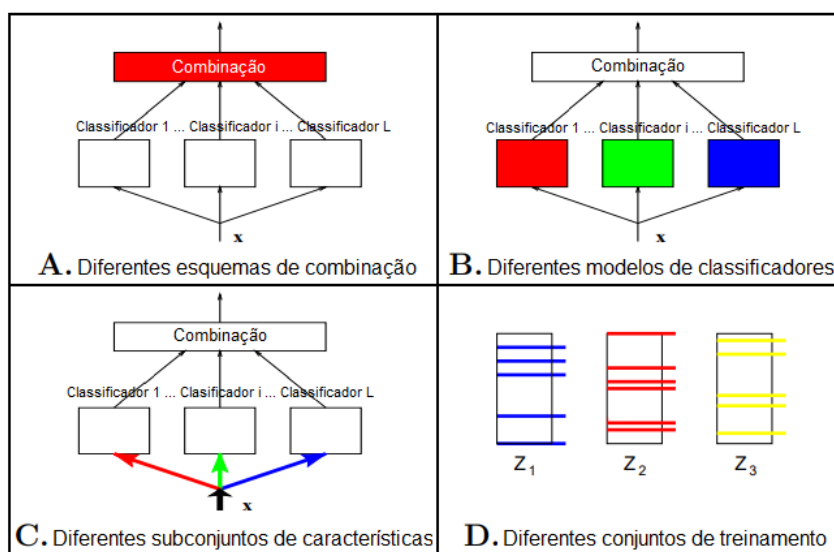
Figura 2.17: Duas principais arquiteturas de *ensemble*. S_i , h_i e $H(x)$ indicam a i -ésima amostra de dados, o i -ésimo classificador e a regra de combinação, respectivamente.



Fonte: Adaptado de (PONTI JR., 2011).

- (A) **Ao nível de combinação:** Assumindo os classificadores, previamente treinados, o problema é escolher um esquema de combinação entre eles e treiná-lo, se necessário;
- (B) **Ao nível de classificadores:** Qualquer classificador pode ser utilizado no *ensemble*. A composição pode ser homogênea, um modelo constituído por classificadores idênticos, ou heterogênea, quando os classificadores são distintos;
- (C) **Ao nível de características:** Treinar um conjunto de classificadores usando diferentes subespaços do espaço de características. Isso é útil quando existe um número expressivo de amostras e os grupos de características têm origens distintas;
- (D) **Ao nível de dados de treinamento:** Treinar um *ensemble* de classificadores usando diferentes conjuntos de treinamentos por divisão (PONTI; PAPA, 2011), por *cross-validation ensembles* (PARMANTO; MUNRO; DOYLE, 1996), *bagging* (??) e *boosting* (FREUND; SCHAPIRE, 1996).

Figura 2.18: Ilustração das quatro abordagens para a criação de um *ensemble*.



Fonte: Adaptado de (KUNCHEVA, 2001).

Além disso, há várias possibilidades para a combinação dos classificadores. As duas principais abordagens são por integração (ou fusão) e seleção. Na primeira, todos os classificadores contribuem para a decisão final, assumindo que os classificadores são competitivos. Na segunda, um classificador é usado para dar a decisão final para cada padrão, assumindo que os classificadores são complementares. Quanto ao tipo de informação produzido pelos classificadores do *ensemble*, três níveis são possíveis: abstrato, cada classificador produz o rótulo de classe para cada padrão de entrada; *rank*, cada classificador produz como saída um ranqueamento de possíveis classes para cada padrão de entrada; e medida, cada classificador produz uma pontuação, probabilidade ou nível de confiança para cada padrão de entrada. Assim, considerando as formas de combinações dos resultados dos classificadores, as regras fixas são utilizadas quando as saídas estão ao nível de medidas (NANNI et al., 2020; ROBERTO et al., 2021; XIA; ZONG; LI, 2011), podendo assumir diferentes regras (PONTI JR., 2011):

- **Produto:** Multiplica as pontuações providas por cada classificador (em cada classe) e atribui o rótulo da classe que possui a pontuação máxima para determinado padrão de entrada;

- **Soma:** Soma a pontuação fornecida por cada classificador (em cada classe) e atribui o rótulo da classe que possui a pontuação máxima para determinado padrão de entrada;
- **Média:** Calcula a média das pontuações de cada classe entre os classificadores e atribui o padrão de entrada à classe que possui a maior pontuação entre as médias calculadas. É equivalente à regra da soma.

Na Figura 2.19 é apresentado um problema de classificação composto por quatro classes, denominadas por ω , em um *ensemble* de três classificadores, cada classificador indicado por h . Cada célula da tabela ilustrada na Figura 2.19 relaciona o classificador com uma classe e informa a pontuação que o mesmo obteve (linha) na classe (coluna). Nessa tabela estão indicadas as pontuações de cada classe, considerando o *ensemble* baseado na regra fixa descrita. A classe resultante é indicado à extrema-direita e o resultado responsável por essa decisão está indicado em cinza.

Além das regras fixas baseadas em medidas, descritas anteriormente, na Figura 2.19 também é apresentada a estratégia de combinação baseada em voto majoritário. Essa abordagem trabalha ao nível abstrato, ou seja, produz como saída um rótulo referente à classe. Para poder combinar os classificadores, o método de votação encontra qual é a classe de saída de cada classificador e contabiliza essa saída como um voto para a classe. O padrão de entrada é atribuído à classe com o voto majoritário, aquele que possui mais votos. Além do voto majoritário, existem outras variantes, como a votação ponderada. Para evitar o problema de empate, são normalmente utilizados um número ímpar de classificadores (PONTI JR., 2011).

2.11 Métricas para Avaliação de Desempenho

Os resultados produzidos pelos classificadores são contagens de classificações corretas e incorretas para cada classe (BRADLEY, 1997). Supondo um problema de clas-

Figura 2.19: Combinação baseada em regras fixas em um problema com quatro classes e um *ensemble* de três classificadores.

	ω_1	ω_2	ω_3	ω_4
h_1	0.1	0.5	0.2	0.1
h_2	0.3	0.3	0.3	0.1
h_3	0.2	0.0	0.8	0.0

Regra					Resultado
Mínimo	0.1	0.0	0.2	0.0	ω_3
Máximo	0.3	0.5	0.8	0.1	ω_3
Produto	0.01	0.00	0.05	0.00	ω_3
Média	0.2	0.3	0.4	0.1	ω_3
Mediana	0.2	0.3	0.3	0.1	empate
Voto Majoritário	1	2	2	0	empate

Fonte: Adaptado de (PONTI JR., 2011).

sificação que considera apenas duas classes, cada instância pode ser mapeada a um elemento do conjunto $\{\mathbf{p}, \mathbf{n}\}$, que rotulam as classes positiva e negativa. Para diferenciar a classe atual da predita, os rótulos $\{\mathbf{v}, \mathbf{f}\}$ são utilizados para indicarem a classe predita pelo classificador (FAWCETT, 2006). Dado um classificador e uma instância, quatro rotulações são possíveis (SOKOLOVA; LAPALME, 2009):

- **Verdadeiro Positivo (vp).** Quando uma instância é positiva e classificada positiva, ou seja, a instância é classificada corretamente e pertence à classe em análise;
- **Verdadeiro Negativo (vn).** Quando uma instância é negativa e classificada como negativa, ou seja, a instância é classificada corretamente e não pertence à classe em análise;
- **Falso Positivo (fp).** Quando uma instância é negativa e classificada como posi-

tiva, ou seja, a instância foi atribuída incorretamente à classe em análise;

- **Falso Negativo (fn).** Quando uma instância é positiva e classificada como negativa, ou seja, a instância não foi reconhecida como pertencente à classe em análise.

Quando considerado um classificador e as instâncias do conjunto de teste, é possível construir uma matriz de tamanho 2×2 , denominada matriz de confusão (Figura 2.1), que indica a disposição dessas instâncias de acordo com suas quatro possíveis classificações. Os números na diagonal principal indicam as classificações corretas, enquanto os da diagonal secundária informam as classificações erradas.

Tabela 2.1: Matriz de confusão com as quatro classificações possíveis.

		Classe atual/esperada	
		p	n
Classe Prevista	v	verdadeiro positivo	falso positivo
	f	falso negativo	verdadeiro negativo

Fonte: Adaptado de (FAWCETT, 2006).

A matriz é a base para o cálculo de diversas métricas para avaliação de desempenho do classificador, tais como acurácia e área sob a curva ROC (*AUC – Area Under the ROC Curve*). A acurácia quantifica a efetividade geral do classificador, calculada pela divisão entre a soma das classificações corretas, positivas e negativas, pela soma de todas as classificações:

$$acuracia = \frac{vp + vn}{vp + vn + fp + fn}. \quad (2.6)$$

A AUC é obtida a partir do gráfico da curva ROC (*Receiver Operating Characteristic*) (BRADLEY, 1997), expresso em termos de duas medidas: sensibilidade (Equação 2.7) e especificidade (Equação 2.8). A sensibilidade indica a taxa de amostras positivas que foram classificadas corretamente como positivas. E a especificidade indica a

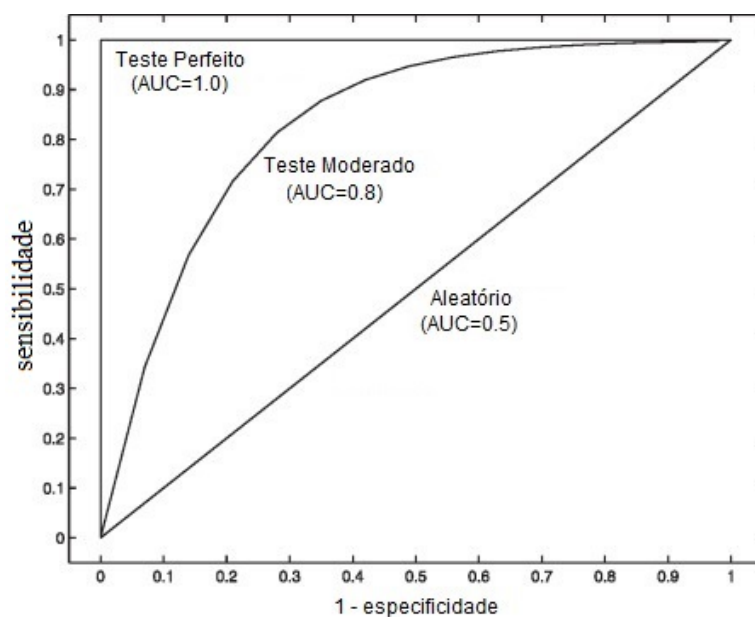
taxa de amostras negativas que foram classificadas corretamente como negativas.

$$sensibilidade = \frac{vp}{vp + fn}. \quad (2.7)$$

$$especificidade = \frac{vn}{fp + vn}. \quad (2.8)$$

O eixo x do gráfico ROC é definido por $1 - \text{especificidade}$ e o eixo y é definido pela sensibilidade, conforme ilustrado na Figura 2.20. A AUC pode assumir valores entre 0 e 1. Uma curva ROC com área 1 é indicado na Figura 2.20, que define uma classificação perfeita. Uma curva que segue em linha reta do ponto (0,0) até o ponto (1,1) possui uma AUC de 0,5. Nessa situação, infere-se que o classificador possui poder de classificação aleatório, uma vez que sua taxa de acerto é de 50%. Valores de área sob a curva ROC maiores de 0,5 indicam classificadores com alguma habilidade de discriminação. Quanto mais próximo a AUC chegar de 1,0, melhor será a discriminação. Um exemplo é a curva com $AUC = 0,8$, na Figura 2.20 (DEVOS et al., 2007).

Figura 2.20: Gráfico de comparação entre três curvas ROC.



Fonte: Adaptado de (DEVOS et al., 2007).

2.12 Imagens Histológicas

Segundo (GARTNER, 2020), histologia refere-se ao estudo microscópico dos tecidos de organismos vivos, sejam eles animais ou plantas. Além disso, a histologia envolve mais do que apenas a estrutura microscópica do corpo, mas também se relaciona ao funcionamento do corpo. O estudo é realizado por meio de microscópio óptico, no qual a imagem se forma a partir dos raios luminosos de um feixe de luz que atravessa uma estrutura (JUNQUEIRA; CARNEIRO, 2013).

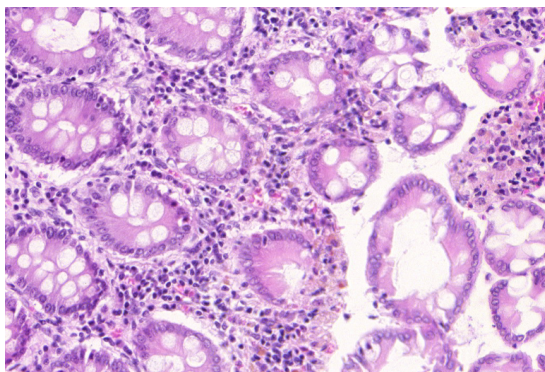
A preparação de uma amostra histológica demanda uma sequência de passos, dentre eles está o processo de coloração dos tecidos. Como muitos constituintes do tecido têm aproximadamente as mesmas densidades, eles devem ser corados para microscopia de luz (GARTNER, 2020). A coloração comumente utilizada na histologia é a Hematoxilina-Eosina (H&E). A hematoxilina é uma base que se liga preferencialmente aos componentes ácidos da célula, como o DNA e RNA, colorindo-os de azul. A eosina é um ácido que se liga aos constituintes citoplasmáticos que possuem um pH básico, corando-os em rosa. Exemplos de imagens histológicas H&E são apresentadas na Figura 2.21.

2.13 Considerações Finais sobre a Fundamentação Teórica

Todos os conhecimentos necessários para o entendimento do trabalho foram explorados neste capítulo. Iniciando pelos conceitos de extração de atributos, foram apresentadas as técnicas fractais, que desempenham importante função na quantificação de texturas e na análise de objetos com a propriedade de autossimilaridade, comumente observadas nas formas da natureza. Diferentes abordagens foram encontradas na literatura, cada uma com a sua particularidade, mas ainda assim complementando umas às outras. Ainda no contexto dos descritores de textura, foram apresentadas as téc-

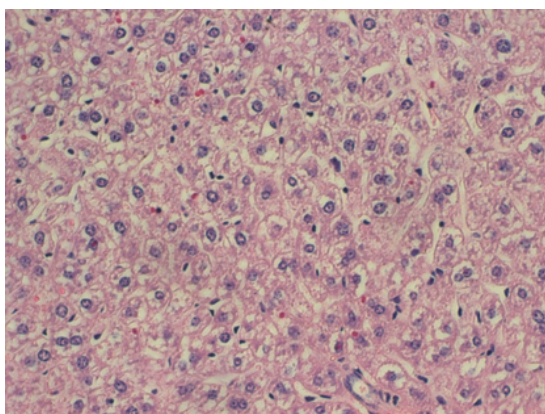
Figura 2.21: Exemplos de imagens histológicas H&E.

(a) Câncer Colorretal – Benigno



Fonte: (SIRINUKUNWATTANA et al., 2017).

(b) Tecido Hepático (LA) – Classe 1



Fonte: (AGEMAP, 2020).

nicas mais clássicas LBP e Haralick, amplamente utilizadas e consolidadas entre os pesquisadores. A segunda categoria de atributos são aquelas obtidas via arquiteturas de redes neurais convolucionais. Para isso, as imagens são dadas como entradas para esses modelos, que processam cada *pixel* das imagens em um processo de convolução, percorrendo as diferentes camadas que compõe uma arquitetura. Os inúmeros valores extraídos dos métodos mencionados passam pelas técnicas de seleção de atributos, cujo intuito é a redução da dimensionalidade do vetor de entrada e, possivelmente, uma melhora no desempenho do sistema. Dentre essas técnicas, foram apresentadas as categorias *wrapper*, baseadas em meta-heurísticas, e de filtragem, com o algoritmo ReliefF. Por fim, o conjunto de atributos passa por um processo de classificação. Nisso,

é apresentado o conceito do *ensemble learning*, que utiliza a estratégia de unir diferentes classificadores para criar um processo de classificação mais robusto, o qual é comumente avaliado por meio de métricas de desempenho. Por fim, foram apresentados os principais conceitos envolvidos no contexto de imagens histológicas.

Capítulo 3

Trabalhos Relacionados

Este capítulo possui como objetivo a apresentação de estudos encontrados na literatura envolvendo as técnicas utilizadas neste trabalho. Mais especificamente, nas Seções de 3.1 a 3.4 são apresentados os trabalhos envolvendo as técnicas utilizadas para extração de características, respectivamente: técnicas fractais multidimensionais e multiescala, descritores de Haralick, LBP e redes neurais convolucionais. Na Seção 3.5 são abordados trabalhos que utilizam as CNNs exclusivamente como extratores de características e que associam esses atributos convolucionais às características *hand-crafted*, em um processo de classificação sem a abordagem de *ensemble*. A Seção 3.6 tem como foco os estudos envolvendo o *ensemble learning* entre as duas categorias de características. Adicionalmente, na Seção 3.7 estão alguns trabalhos correlatos sobre a estratégia de duplo estágio. Por fim, na Seção 3.8 é apresentada uma visão geral dos trabalhos apresentados, com o foco na técnica de *ensemble* no contexto de classificação de imagens médicas.

3.1 Técnicas Fractais Multidimensionais e Multiescala

Os autores Ivanovici, Richard e Decean (2009) aplicaram os métodos de dimensão fractal e lacunaridade multiescala, propostos com base no algoritmo probabilístico

definido por Mandelbrot (1983). O intuito foi a análise de imagens de lesões na pele, mais especificamente do tipo psoríase. Os autores concluíram que peles saudáveis apresentaram valor de dimensão fractal normalmente abaixo de 3,0, mais especificamente em um intervalo entre 2,5 e 2,8, dependendo da complexidade da textura da pele. Além disso, associa-se um baixo valor de lacunaridade às peles saudáveis. Para maior grau de comparação das técnicas, foram realizadas comparações entre os valores de dimensão fractal e lacunaridade com pontuações (grau de severidade) avaliadas pelo especialista dermatológico, considerando três características das lesões psoriáticas: o eritema (vermelhidão), a escamação e o endurecimento (espessura). Disso, concluiu-se que há relação direta entre a gravidade das lesões com a dimensão fractal.

Os autores Roberto et al. (2019) aplicaram a abordagem de percolação multidimensional e multiescala proposta em (ROBERTO et al., 2017) para a análise de imagens de tumores de mama e colorretais normalizadas. O intuito foi o estudo da influência de diferentes técnicas de normalização de cores sobre características baseadas na percolação. Inicialmente, foram extraídas as características das duas bases de imagens sem a normalização, resultando para cada contexto, melhores AUCs maiores do que 0,900. Três tipos de normalização foram aplicados, sendo possível concluir que a normalização da cor é uma abordagem adequada para classificação dos tumores colorretais com as características de percolação. Para o conjunto de dados de câncer de mama, nenhum dos métodos de normalização testados conseguiu fornecer melhores acurácias em relação às imagens não normalizadas, cujo valor foi de 87,93%. Analisando ainda o mesmo conjunto de imagens, houve uma melhora na taxa de AUC, 0,929, com a normalização de cores de Vahadane. Porém, o ganho obtido poderia não ser significativo o suficiente para aplicar a técnica de normalização de cor. Outra comparação foi aplicada: comparar os atributos percolantes com atributos de dimensão fractal multiescala (IVANOVICI; RICHARD, 2011), lacunaridade multiescala (IVANOVICI; RICHARD, 2009), descritores de Haralick (HARALICK; SHANMUGAM; DINSTEN,

1973) e características LBP nas imagens normalizadas e não normalizadas. No conjunto de imagens de mama, a melhor AUC foi obtida pelos descritores de Haralick, $AUC = 0,880$, em imagens normalizadas pelo método Vahadane, seguida pela AUC obtida pelos atributos percolantes, $AUC = 0,879$, em imagens não normalizadas. Para o conjunto colorretal, as melhores taxas de AUC, nas imagens não normalizadas e normalizadas pelos três métodos, foram alcançadas pelo método de percolação multidimensional e multiescala. Os autores concluíram que a aplicação da normalização de cores na base de mama não teve contribuições relevantes quando considerado o método de percolação para extração das características. Por outro lado, para a base de colorretal, quando associada a normalização de cores com os atributos percolantes, houve melhora nos resultados para a classificação das imagens. Por fim, as características fractais globais tiveram papel importante na classificação dos dados colorretais, mas não para os de mama.

Ribeiro et al. (2019) propuseram um método que associa técnicas fractais multidimensionais, transformada *curvelet* e descritores de Haralick para a análise e classificação de imagens de câncer colorretal, cuja base de imagens é composta por 151 imagens histológicas H&E subdivididas nos grupos benigno e maligno. Dentre as técnicas fractais, foram abordadas a dimensão fractal, lacunaridade e percolação. O método analisa as imagens coloridas por meio das técnicas fractais multidimensionais e multiescala e descritores de Haralick. As sub-imagens obtidas pela aplicação da transformada *curvelet* são submetidas da mesma forma aos descritores de Haralick e técnicas fractais, no entanto, para a análise de imagens em níveis de cinza. As características obtidas são concatenadas em um vetor e aplicadas ao método de seleção ReliefF. Por fim, seis classificadores foram utilizados para predizer a qual classe uma imagem pertence: ao grupo benigno ou ao grupo maligno. A melhor taxa de AUC obtida foi de 0,994 com o classificador polinomial.

3.2 Descritores de Haralick

Os autores Wibmer et al. (2015) propuseram uma análise dos descritores de Haralick na análise de imagens de ressonância magnética (MRI) para a detecção do câncer de próstata e na diferenciação dos escores de Gleason (GS). Esses escores são pontuações dadas a um câncer de próstata baseado em sua aparência microscópica. Cento e quarenta e sete pacientes foram submetidos à ressonância magnética por difusão (DWI) e ponderada em T2 (T2WI). Câncer e não-câncer, subdivididas em periférica (PZ) e transição (TZ), foram identificados nas imagens T2WI e nos mapas de coeficiente de difusão aparente (ADC) — utilizados na análise quantitativa da difusão. Os autores aplicaram cinco descritores de Haralick: energia, entropia, correlação, homogeneidade (momento inverso da diferença) e contraste, analisadas usando equações de estimativa generalizadas. O uso de apenas cinco características é justificado com base no trabalho (CONNERS; TRIVEDI; HARLOW, 1984), que eliminou características redundantes e reduziu o número de atributos necessários. Como resultado, o câncer PZ apresentou maior entropia e inércia e menor energia, correlação e homogeneidade comparado aos tecidos não-câncer nos T2WI e nos mapas ADC. No câncer TZ, foram observadas diferenças significantes entre os cinco descritores no mapa ADC e para correlação e inércia considerando T2WI. Nos mapas ADC, o valor de GS foi associado a uma alta entropia e baixa energia. Foi possível concluir que as características da textura de Haralick derivadas de imagens ponderadas em T2 e de mapas ADC têm o potencial de diferenciação entre câncer de próstata e tecido não-canceroso da próstata, além da avaliação do *score* de Gleason.

Em um estudo sobre os descritores de Haralick, os autores Caruso et al. (2018) propuseram a investigação de quais descritores seriam mais viáveis na resposta do tumor à quimiorradioterapia neoadjuvante (CRT) nos casos de câncer colorretal. Para o estudo, foram obtidas imagens MRI de pacientes que possuíam o adenocarcinoma colorretal comprovado histologicamente e cujo tumor estava localmente avançado. Esses paci-

entes foram divididos em duas classes baseados no critério de responder totalmente ao tratamento (CR) ou não responder (NR). Os resultados preliminares mostraram que 5 dos 14 descritores apresentaram relevância significativa na previsão da resposta à terapia nos dois grupos de pacientes em análise. Os descritores identificados foram: energia, fornecendo conhecimento sobre uniformidade de imagem; contraste, medindo a variação de intensidade entre os *pixels* da imagem; correlação, reconhecendo o paralelismo entre os níveis de cinza da imagem, indicando as dependências lineares; entropia, indicando a distribuição caótica entre os níveis de cinza; e momento inverso das diferenças, se referindo à uniformidade que quantifica a afinidade dos níveis de cinza da co-ocorrência. Além desses resultados, os autores indicam a associação desses descritores com dados genômicos adicionais, que poderiam fornecer uma visão mais abrangente das características do tumor, possibilitando uma terapia mais direcionada e característica ao paciente em questão.

Os autores Fondón et al. (2018) abordaram a detecção automática da malignidade de amostras de tecidos mamários com base no processamento de imagens histológicas H&E. São quatro classes de saída: normal, benigno, *in situ* e invasivo. O método se baseia na extração de três conjuntos de características: aquelas relacionadas ao núcleo, às regiões coloridas e à textura (considerando características locais e globais). Quanto a textura, são extraídos os descritores a partir do método LBP e pelos métodos “clássicos”, sendo eles: estatísticas de primeira ordem, dimensão fractal global (dimensão fractal de Hausdorff) e os 14 descritores de Haralick. Anteriormente à fase de extração, realiza-se um pré-processamento para a normalização de cor das imagens. Os testes foram realizados seguindo três diferentes conjuntos de imagens. Na primeira fase de testes considerou-se um total de 9 classificadores, testados em um conjunto composto por 120 imagens, sendo 30 imagens de cada classe que tendem a serem o mais representativo possível segundo a ampla variedade de aparências que podem ser encontradas em imagens H&E. O melhor desempenho foi obtido pelo classificador SVM

com *kernel* quadrático, cuja acurácia foi de 75,80%. No segundo conjunto, 20 imagens de mesma complexidade são utilizadas para validar externamente o método. Para isso, o classificador com melhor acurácia no primeiro teste foi submetido ao novo conjunto de imagens, resultando em uma acurácia de 75,00%. Por fim, o terceiro conjunto de imagens foi composto pela união do conjunto do segundo teste com 16 imagens extremamente difíceis de serem analisadas, pois a classificação do nível de malignidade é mais ambígua. A acurácia obtida foi de 61,11%. O método também foi comparado a outro trabalho disponível na literatura baseado no uso de CNN, apresentando resultados bem próximos quando aplicados nos conjuntos de imagens 2 e 3.

3.3 Descritores LBP

No trabalho dos autores Mazo, Alegre e Trujillo (2017), os descritores LBP e sua variante LBPri foram explorados na classificação de imagens histológicas de tecidos cardiovasculares. Para isso, inicialmente, foi realizada a extração dos atributos a partir de blocos não sobrepostos de dimensões 100×100 , obtidos das imagens originais. Foram explorados um total de 3.000 blocos/imagens, divididos em 600 blocos por classe analisada. Essas classes foram divididas em quatro grupos: (1) coração, músculo cardíaco do coração; (2) muscular, músculo liso da artéria muscular; (3) conectivo, tecido conjuntivo frouxo; (4) elástico-veia; sendo esse último grupo subdividido em: (4.1) elástico, músculo liso da artéria elástica; (4.2) veia, músculo liso da grande veia. Os vetores de características obtidos da etapa anterior foram entradas para o algoritmo de classificação SVM com *kernel* linear. O algoritmo classificou as imagens de entrada nos grupos de (1) a (4). Um segundo passo de classificação foi realizado. O classificador SVM com *kernel* polinomial foi aplicado para separar a classe (4) nas classes (4.1) e (4.2). Os valores de AUC obtidos na fase de classificação foram superiores a 0,98 em todas as classes analisadas.

No contexto de imagens histológicas de câncer de mama, os autores Yu et al.

(2019) propuseram um método de detecção automática baseado em características híbridas para a classificação das imagens nos grupos maligno e benigno. A metodologia explorou as imagens manipuladas de duas formas: primeira, segmentadas em três classes por uma CNN, divididas em núcleo, limite nuclear e fundo. Posteriormente as imagens foram pós-processadas, resultando nas imagens finais dos núcleos segmentados; segundo, manipulação das imagens de entrada envolvendo os canais de cores H (hematoxilina) e E (eosina). Das imagens com segmentação de núcleos foram extraídos atributos morfológicos e espaciais. Das imagens com a decomposição dos canais de cores, foram extraídas as características de textura, como descritores de Haralick e LBP. O vetor de características foi dado como entrada ao algoritmo Relief, que ranqueou os 50 melhores atributos. Em seguida, foi aplicado o classificador SVM. As taxas obtidas foram uma AUC de 0,983 e uma acurácia de 96,70%, apresentando desempenho superior em comparação com outras técnicas existentes.

3.4 Redes Neurais Convolucionais

Um estudo foi proposto por Spanhol et al. (2017) para analisar o desempenho do classificador treinado a partir de características extraídas de uma CNN treinada no banco de imagens ImageNet. A arquitetura utilizada foi a CaffeNet, que consiste em uma pequena modificação da arquitetura AlexNet. Dada a variabilidade de classes no qual a rede foi treinada, supõe-se inicialmente que a representação aprendida com uma CNN treinada neste conjunto de dados define um extrator de características de uso geral ótimo. As camadas das quais se extraiu as características da CaffeNet (características DeCAF) foram as camadas fc6, fc7 e fc8. Para testar o desempenho das DeCAF, o classificador de Regressão Logística foi testado no banco de imagens específico (BreakHis), dando como entrada cada camada separadamente e as quatro combinações possíveis entre elas. O pior conjunto de características individuais foi obtido pela camada fc8. No geral, embora seja possível observar algumas melhorias na acurácia para

alguns casos considerando a combinação de características, a maior margem de ganho, em comparação com os melhores resultados obtidos com uma única camada individualmente, foi de apenas 0,3%, sendo um aumento de 86,00% para 86,30% considerando um fator de ampliação $200\times$ das imagens e, 84,30% para 84,60% com amplificação $40\times$. Outro tipo de comparação foi realizado: considerando o método proposto em relação às características *handcrafted* tradicionais e uma CNN treinada em um conjunto de imagens específica ao problema. Como métrica de comparação, utilizou-se o *F1 score*. O método se sobressaiu em relação às características *handcrafted* e obteve resultados semelhantes quando comparado à CNN treinada para a tarefa específica. A melhor *F1 score* obtido pelo método foi de 90,30% contra 93,50% obtida pela CNN usando combinação de múltiplos classificadores e 90,00% sem essa combinação. Esses resultados se demonstraram importantes para o projeto de futuros sistemas baseados em sistemas CADs, uma vez que mostra que características profundamente aprendidas, mesmo que obtidas com uma CNN treinada em outros tipos de imagens, são valiosas.

Araújo et al. (2017) utilizaram a arquitetura de CNN para a classificação de imagens histológicas H&E de câncer de mama. Além da classificação binária entre carcinoma e não-carcinoma, quatro classes foram analisadas: tecido normal, lesão benigna, carcinoma *in situ* e carcinoma invasivo. O conjunto de imagens é composto por 249 imagens para treinamento e 20 imagens para teste. Como o número de amostras é baixo para realizar o treinamento da rede, utilizou-se a estratégia de dividir as imagens em *patches* e criar amostras a partir da técnica de *data augmentation*. As imagens foram classificadas segundo duas abordagens: a primeira nomeada de *patch-wise classification*, que consiste em classificar as imagens utilizando a própria CNN e também um classificador SVM, treinado a partir das características extraídas da CNN; a segunda abordagem é a *image-wise classification*, em que a classificação da imagem é realizada primeiro processando vários *patches* segundo a abordagem *patch-wise clas-*

sification, e após, combinando os resultados da classificação de todos os *patches* da imagem para obter a classificação final da imagem. Os resultados obtidos no primeiro método, considerando as quatro classes, foi uma acurácia de 66,70% pela CNN somente e, 65,00% considerando o classificador SVM treinado com as características obtidas pela rede. Em relação à classificação binária, houve um aumento da acurácia, de modo que a CNN produziu acurácia de 77,60% e 76,90% com CNN+SVM. Na segunda abordagem de classificação, a acurácia foi de 77,80% na classificação das quatro classes, considerando tanto a CNN como CNN+SVM na fase de *patch-wise classification*. Por fim, foi possível notar melhora na acurácia para a classificação binária quando considerado o modelo CNN+SVM, cuja acurácia foi de 83,30%. Com esses resultados, os autores concluíram que o método proposto permite obter alta sensibilidade para os casos de carcinoma e, embora um conjunto de dados menor e mais desafiador tenha sido utilizado, o desempenho do sistema é semelhante ou superior a outros métodos disponíveis na literatura.

Os autores Rezvantab, Safigholi e Karimijeshni (2018) propuseram um estudo sobre a eficácia e capacidade das redes neurais convolucionais na classificação de câncer de pele divididos em oito categorias. O objetivo é comparar a capacidade de aprendizado profundo com o desempenho de dermatologistas altamente treinados. Foram utilizados dois bancos de imagens: um contendo 10.015 imagens de sete categorias de diagnósticos; o segundo banco é composto por 200 imagens, dos quais foram utilizadas duas categorias, sendo uma delas distintas das sete da base anterior, somando as oito classes. O método consistiu em utilizar quatro CNNs (DenseNet-201, ResNet-152, InceptionV3 e InceptionResNetV2) pré-treinadas no banco de imagens ImageNet. Cada uma dessas redes passaram pelo processo de *fine-tuning*. Como o conjunto de dados é desequilibrado, foram calculadas médias macro e micro das taxas de AUC para as comparações. A arquitetura que obteve o melhor desempenho considerando as oito classes foi a DenseNet-201, com média de AUC macro sendo 98,16% e micro

de 98,79%. A rede que obteve os piores resultados foi a InceptionResNetV2, com AUC 98,20% (micro) e 96,10% (macro). Para realizar a comparação considerando a avaliação dos dermatologistas, a AUC considerada nesse caso foi a média das taxas de sensibilidade e especificidade. Dada a classe melanoma, a melhor e a pior AUC obtidas pelas CNNs foram de 94,40% com a ResNet-152 e de 93,20% com a InceptionResNet, respectivamente, enquanto a AUC obtida pelo dermatologista foi de 88,82%. De maneira similar, considerando o tipo carcinoma basocelular, a melhor AUC foi de 99,30% com a rede DenseNet-201 e de 88,82% pelo dermatologista. Concluindo que com o uso de CNNs e *fine-tuning* em imagens dermatoscópicas, pode-se obter uma melhor precisão diagnóstica em comparação com médicos e clínicos especialistas.

Ghosh et al. (2021) utilizaram arquiteturas CNN para comporem um modelo de *ensemble*, a fim de detectar tumores colorretais em imagens histológicas. O *ensemble* trabalhou com dois conceitos: um modelo sequencial e um paralelo. A técnica sequencial foi aplicada progressivamente avaliando a dependência entre os modelos-base, as próprias CNNs. No paralelo, os modelos foram gerados simultaneamente e a justificativa, neste caso, foi utilizar a dependência entre os modelos. Com isso, as arquiteturas DenseNet-121, InceptionResNetV2, Xception e uma quarta arquitetura customizada pelos autores foram treinadas separadamente, segundo o modelo paralelo. Posteriormente, foram atribuídos os dados de entrada para o modelo de *ensemble* sequencial, nos quais as melhores características (pesos) de cada rede foram extraídas, baseadas em métricas como acurácia e precisão, e implementadas como entrada no modelo. O *ensemble* foi treinado congelando todas as camadas, exceto as camadas totalmente conectadas, e validado. Duas bases de dados foram utilizadas: uma contendo 5.000 imagens e outra com 107.180 imagens. Na primeira base, os autores relataram que a acurácia obtida foi de 92,83%. Na segunda base, a acurácia foi de 96,16%. Por fim, quando os bancos de imagens foram combinados, a acurácia aumentou para 99,13%.

Ohata et al. (2021) elaboraram uma abordagem voltada para a detecção automá-

tica de oito tipos de tecidos histológicos de câncer colorretal. Para tal, fizeram o uso de dezoito arquiteturas CNNs pré-treinadas no *dataset* ImageNet, dentre elas as famílias da VGG, Inception, ResNet e DenseNet. As camadas totalmente conectadas foram extraídas, mantendo as camadas convolucionais e de *pooling*, de forma que a saída resultante das redes formaram, separadamente, a entrada para os classificadores. Foram considerados os algoritmos de classificação Bayes, *Multilayer Perceptron*, KNN, *Random Forest* e SVM. Cada conjunto de *deep learned features* foram treinados separadamente com cada classificador, resultando em 108 combinações. A melhor combinação obtida foi com a arquitetura DenseNet-169 e classificador SVM (*kernel* RBF, cuja acurácia atingiu 92,08%).

3.5 *Ensemble de Descritores Deep Learned e Handcrafted sem o uso Ensemble de Classificadores*

Os autores Almaraz-Damian et al. (2020) propuseram um sistema para a detecção e classificação de lesões na pele, subdividido nos grupos nevus e melanoma. Para isso, os autores utilizaram características *handcrafted* obtidas a partir da aplicação dos descritores de Haralick e do algoritmo médico ABCD (*Asymmetry- Borders-Colors-Dermatoscopic Structures*), que se baseia na cor, forma e estruturas particulares observadas em lesões na pele. Além disso, foram obtidos descritores de forma, como área, perímetro, circularidade, diâmetro e excentricidade das regiões de interesse das imagens previamente segmentadas. Diferentes arquiteturas de CNNs previamente treinadas no banco de imagem *ImageNet* foram aplicadas como extratores de características. As características *deep learned* foram combinadas com as características *handcrafted*, separadamente, e aplicadas ao algoritmo de informação mútua. Os classificadores utilizados foram regressão logística, SVM e *Relevant Vector Machine*. O melhor resultado apresentado foi a partir da fusão das características *handcrafted* com as carac-

terísticas extraídas da arquitetura Mobilenet v2, cuja acurácia foi de 92,40%. Quando comparado a outros trabalhos na literatura, apresentou resultados superiores.

Com o intuito de desenvolver uma arquitetura de *deep learning* para apoiar na detecção automática de câncer no cérebro a partir de imagens de ressonância magnética, Rajinikanth et al. (2020) propuseram um modelo baseado na fusão das características extraídas da arquitetura VGG-19 e de características obtidas pelas transformadas *wavelet*, *curvelet* e *contourlet*. Para isso, os autores aplicaram a técnica de transferência de aprendizado, testaram e validaram com o classificador *SoftMax* diversas arquiteturas de CNNs, tais como AlexNet, VGG-16, VGG-19, ResNet-50 e ResNet-101. A rede que ofereceu a melhor acurácia foi escolhida para compor o método, sendo esta a arquitetura VGG-19. O classificador *SoftMax* foi então substituído pelos classificadores árvore de decisão, KNN e SVM com *kernel* linear e RBF (função de base radial). Ao total, foram obtidas 199 características *handcrafted* e 1024 características da arquitetura VGG-19. Esses valores foram concatenados em um vetor, em que foi aplicada uma extração de atributos pelo algoritmo PCA (análise de componentes principais) e, posteriormente, aplicados no treinamento, teste e validação do novo classificador implementado na rede. A fusão de ambas as características é justificado pelo fato de que em pesquisas anteriores, os autores observaram que se o vetor de características for aprimorado, a arquitetura de *deep learning* pré-treinada oferecerá melhor desempenho na predição das classes em estudo, quando comparado com a arquitetura tradicional. Três categorias de imagens foram analisadas: Flair, T2 e T1C, subdivididos em normal e tumor. Os melhores resultados em todos os grupos foram obtidos com o classificador SVM-RBF, que apresentou acurácias de 99,10% (Flair), 98,90% (T2) e 97,70% (T1C).

O estudo proposto por Sethy e Behera (2022) utiliza os atributos *deep learned* e *handcrafted* para a classificação de imagens H&E de câncer de mama em quatro classes. Na primeira categoria de atributos, foram explorados aqueles referentes à camada FC6 nas arquiteturas AlexNet, VGG-16 e VGG-19 e, dentre os *handcrafted*, foram uti-

lizados GLCM, HOG e LBP. Para chegarem ao modelo final, os autores aplicaram três fases de construção. A primeira delas consistiu em classificar, separadamente, as *deep features* utilizando três diferentes classificadores – Naive Bayes, KNN e SVM. Na segunda parte, utilizando a mesma abordagem, foram avaliados os descritores *handcrafted*. Dessas duas etapas, os descritores com melhores desempenhos foram os da rede VGG-16 e LBP. Finalmente, para a terceira e última etapa, os autores propuseram o modelo final que toma como entrada o *ensemble* entre descritores LBP e VGG-16, seguida por um estágio de classificação via *fine*-KNN para categorizar as imagens H&E nas suas respectivas classes. A acurácia e AUC obtidas pelo modelo foram 84,20% e 0,85, respectivamente.

Os autores Hassan et al. (2022) propuseram a utilização do *ensemble* de descritores *handcrafted* e *deep learned* para a classificação de imagens histológicas de câncer de mama subdivididos em quatro classes. O método se inicia com um processo de normalização de cores e *data augmentation* nas 400 imagens do *dataset* ICIAR-2018 *challenge*. Após isso, são extraídos os atributos *handcrafted* – descritores de Haralick, histograma colorido, *Rotation and Scale-invariant Hybrid Image Descriptor* (RSHD), *Local Diagonal Extrema Pattern* (LDEP), *speeded up robust feature* (SURF) e *Dense Scale Invariant Feature Transform* (DSIFT). Além destes, foram explorados os atributos *deep learned* das seguintes arquiteturas nas camadas após a *avg_pool* global: ResNet-18, Inception-ResNet v2, ShuffleNet e Xception. Os autores exploraram diferentes combinações de *ensemble* de atributos – entre *handcrafted*; entre *deep learned*; e entre ambas as categorias. Esses *ensembles* foram dados como entrada no algoritmo PCA, a fim de reduzir a sua dimensionalidade. Para a classificação, foram utilizados os classificadores *Random Forest*, SVM e KNN. Entre as *handcrafted features*, o histograma e o descritor RSHD foram os que performaram substancialmente melhor, com acurácias máximas de 79,67% e 80,65% no classificador *Random Forest*, respectivamente. Entre o *ensemble* de *deep features*, a melhor acurácia foi obtida via *ensemble*

Inception ResNet v2 + ResNet-18 + Xception no mesmo classificador, com acurácia de 94,61%. Entretanto, o resultado do *ensemble* entre *deep learned* e *handcrafted features* se mostraram capazes de atingirem melhores desempenhos em comparação às outras combinações realizadas, com acurácia de 96,97% no classificador *Random Forest*.

3.6 *Ensemble de Descritores Deep Learned e Handcrafted* com o uso de *Ensemble* de Classificadores

Li et al. (2018) descreveram um modelo que incorpora as características extraídas das arquiteturas de redes neurais convolucionais modificadas Dense-201 e ResNet-50, junto a características que representam os critérios clínicos adotados por dermatologistas. O trabalho visou realizar o diagnóstico automático de lesões na pele, subdivididas em sete grupos distintos. A metodologia foi dividida em duas etapas. A primeira etapa caracteriza-se pela obtenção das características. Para isso, as duas arquiteturas foram modificadas para se adaptarem ao número de classes que deverão analisar e, posteriormente, as redes foram treinadas segundo o processo de *fine tuning*. Na prática, existem critérios clínicos utilizados para o diagnóstico desse tipo de doença, auxiliando os dermatologistas a realizarem essa tarefa precisamente. Esses critérios incluem características baseadas em cor, textura e morfologia, e serviram como base para que os autores extraíssem computacionalmente tais atributos das imagens. O segundo passo do método consistiu no *ensemble* das características extraídas pelas redes neurais e pelos métodos *handcrafted*, por meio do algoritmo LightGBM. Foram analisadas um total de 10.015 imagens coloridas no padrão RGB, extraídas do banco de imagens ISIC 2018. Dos testes realizados, analisando cada rede separadamente, o conjunto total das características *handcrafted* e o *ensemble* entre as duas técnicas, observou-se que o *ensemble* obteve melhor acurácia, sendo de 85,50% no conjunto de

treinamento e 85,30% no conjunto de teste.

Nanni, Ghidoni e Brahnham (2018) propuseram um *ensemble* entre CNNs e diferentes descritores *handcrafted* para a classificação de bioimagens coloridas. Para o estudo, foram utilizados onze descritores *handcrafted*, oito arquiteturas CNNs e quatro bases de imagens (linfoma, carcinoma mamário, laringe e colorretal). Para cada descritor foi treinado uma SVM. Dessa forma, foram obtidas as acurácias de cada um deles separadamente e do *ensemble* entre todos, combinados pela regra da soma. Desse teste concluiu-se que o *ensemble* obtinha melhor performance quando comparado aos descritores isoladamente. Em relação às CNNs, foram realizados diferentes combinações de *ensemble* pela regra da soma: entre todos os diferentes CNNs treinadas variando a configuração dos parâmetros de aprendizagem e *batch size*, *ensemble* denominado Fus; entre todas as redes anteriores, porém após o processo de *fine tuning*, conjunto denominado FCN-st; e de todas as redes do primeiro conjunto citado, porém com o processo de *fine tuning* baseado em *Two-Round Tuning*, combinação denominada FCN-Two. Após análise dos resultados, conclui-se que os *ensembles* FCN obtiveram melhor desempenho em relação ao Fus e que o conjunto FCN-Two obteve performance similar em relação ao FCN-st. Tendo esses resultados, outro teste foi realizado a fim de verificar a performance entre os *ensembles* somente entre as CNNs (conjunto denominado de FCN+) e do *ensemble* entre as CNNs e as características *handcrafted*. Nesse último, dois *ensembles* foram definidos: Here1 – composto pelas CNNs e por nove características *handcrafted*; Here2 – composto pelas CNNs e por todas as onze características *handcrafted*. Como resultados, nas bases de imagens de linfoma e colorretal, o *ensemble* entre CNNs e *handcrafted* resultaram no melhor desempenho, mais especificamente no conjunto Here1, com acurácias de 97,33% e 97,26%, respectivamente. Nas imagens de carcinoma mamário, o melhor desempenho (96,33%) foi obtido pelo conjunto FCN+, em contraposição com Here1 e Here2, que obtiveram o mesmo desempenho de 95,33%. Na avaliação da base de linfomas, a me-

lhor acurácia foi obtida pelo conjunto Here2 (95,45%), seguido por Here1 (95,38%). Com isso, os autores concluíram que os *ensembles* Here1 e Here2 obtiveram melhores desempenhos quando comparado com FCN+ e, uma vez que Here1 é mais simples de se calcular, obtendo resultados semelhantes ao Here2, recomenda-se o uso do ensemble Here1.

Os autores Hagerty et al. (2019) utilizaram o método de *ensemble* entre características *handcrafted* e *deep learned* para a classificação binária de imagens dermatoscópica de melanoma. Sete descritores foram obtidos das imagens e utilizados para criarem individualmente sete modelos de regressão logística. Além disso, utilizou-se a arquitetura ResNet-50 pré-treinada pelo processo de transferência de aprendizado, no qual foram extraídos 2048 atributos convolucionais. O total de atributos *deep learned* foi reduzido para os 1024 mais relevantes, que serviram para treinar a oitava regressão logística. O *ensemble* entre os oito modelos foi realizado e testado em dois bancos de imagens, um composto por 9174 imagens (HAM10000) e outro por 1636 (NIH SBIR). O resultado obtido pela regressão linear no conjunto de atributos convolucionais no banco de imagens NIH foi um valor de AUC de 0,83. Quando considerado o resultado do *ensemble* entre as características *handcrafted*, a AUC foi de 0,90. No entanto, foi possível notar uma melhor performance, AUC de 0,94, para o *ensemble* entre os dois modelos. Em relação ao banco de imagens HAM10000, a AUC com as características *handcrafted*, *deep learned* e o *ensemble* foi de 0,86, 0,88 e 0,90, respectivamente.

Os autores Nanni et al. (2020) propuseram o *ensemble* de características *handcrafted* com as características extraídas de diferentes arquiteturas de CNNs. O intuito do trabalho foi a classificação de células cervicais. As características *handcrafted* foram obtidas a partir de doze técnicas distintas. Em relação às CNNs, nove modelos foram escolhidos. Quanto ao treinamento dessas redes, utilizou-se a técnica de *fine tuning*, ao invés de iniciar o treinamento do zero. Para cada descritor, um classifica-

dor diferente foi treinado. Para cada camada dos quais foram extraídos os atributos convolucionais, treinou-se uma SVM distinta e, ao fim, os classificadores foram combinados pela regra da soma. Dois testes foram executados: o primeiro deles testa o poder discriminativo das características *handcrafted*, avaliando cada característica separadamente e o *ensemble* de todas elas; no segundo teste, considera-se o *ensemble* dessas características com as *deep learned*. Concluiu-se que o *ensemble* entre as características *handcrafted* obtiveram melhores resultados que considerando apenas uma delas. Quando considerado o conjunto de características obtidas a partir do *ensemble* entre as abordagens *handcrafted* e *deep learning*, no qual a melhor acurácia foi de 94,88%, houve a superação dos resultados apresentados por aqueles que possuem somente as características *deep learned* (94,66%). Ou seja, as características *handcrafted* melhoram a performance do *ensemble* de CNNs. Além do contexto proposto, também foram consideradas bases de imagens de outros contextos para uma validação mais profunda do método proposto, são elas: carcinoma mamário (95,33%), linfoma (97,33%), laringe (95,45%) e colorretal (97,26%), com as melhores acurácias obtidas pelo *ensemble* de características indicadas entre parênteses. Dessa forma, o sistema proposto atinge forte poder discriminativo apesar de não estar ajustado em um conjunto de dados específico, ou seja, os mesmos descritores com os mesmos parâmetros são usados em todos os conjuntos de dados.

Roberto et al. (2021) propuseram um *ensemble* envolvendo abordagens fractais e redes neurais convolucionais para a análise de imagens histológicas de tumores colorretais, linfomas, tumores de mama e tecidos do fígado. O método permitiu extrair as características *handcrafted* por meio de técnicas fractais multidimensionais e multiescala como percolação, lacunaridade e dimensão fractal, considerando três medidas de distância distintas. Ao total, 100 características foram obtidas para cada medida de distância e rearranjadas na forma de uma matriz 10×10 , compondo os canais de cores R, G e B, resultando em uma imagem colorida ($10 \times 10 \times 3$). As imagens geradas

nesse processo foram dadas como entradas a uma CNN, nomeada como F-CNN, cuja saída foi um vetor que indica a probabilidade de qual classe a imagem pertence. Outra CNN foi utilizada, porém, recebendo como entrada as imagens originais, nomeada por O-CNN. As probabilidades das classes obtidas pela saída da camada *SoftMax* dessa rede foram somadas às classes de probabilidades de saída da F-CNN. A técnica de transferência de aprendizado foi aplicada utilizando modelos pré-treinados no banco de imagens ImageNet. Para escolher a CNN mais apropriada, foram analisadas quatro modelos (ResNet-50, ResNet-101, InceptionV3 e Xception), considerando o desempenho em questão de tempo e perda. A rede escolhida, ResNet-50, foi a que obteve o menor tempo de treinamento. Como resultados, foram obtidas as seguintes acurácias: 95,55% (linfomas), 99,39% (colorretal), 89,66% (mama) e 99,62% (fígado). O método ainda foi comparado com outras técnicas como: descritores de Haralick, LBP, percolação, lacunaridade e dimensão fractal. O primeiro teste foi realizado comparando essas técnicas e as duas redes sem considerar o *ensemble*. Como resultados, a rede O-CNN obteve as melhores performances, exceto no banco de imagens de mama, cuja acurácia foi de 81,03%, em comparação com o método de percolação, com acurácia de 82,76%. Por fim, no segundo teste a comparação foi realizada considerando o *ensemble* entre as duas arquiteturas, obtendo então, as melhores acurácias em todos os bancos de imagens analisados.

3.7 Duplo estágio de Seleção de Atributos

Zhang et al. (2018) aplicaram uma estratégia de duplo estágio de seleção voltado para monitorar as condições de máquinas rotativas e diagnóstico de possíveis falhas. Para isso, propuseram um esquema de análise envolvendo extração de atributos, pré-seleção e re-seleção de atributos. Os dados consistem em sinais de vibrações que são decompostos e utilizados para a extração de atributos no domínio da frequência. Em seguida, o estágio de pré-seleção de atributos foi aplicado com o uso do algoritmo

ReliefF, selecionando os atributos mais relevantes no espaço dimensional original. Os descritores foram candidatos no estágio de re-seleção, com a aplicação do algoritmo *wrapper* BPSO-SVM (*Binary particle swarm optimization*). A segunda etapa da seleção de atributos (BPSO) foi definida durante a otimização dos parâmetros do classificador SVM. Os autores testaram a proposta em dois casos: (1) diagnóstico de falha de rolamento; (2) diagnóstico de falha do sistema do rotor. As acurácias obtidas foram 98,57% e 99,75%, respectivamente.

Visando a classificação de imagens de radiografias do tórax nas classes COVID-19, pneumonia e normal, os autores Bhowal, Sen e Sarkar (2021) propuseram um método que utiliza as arquiteturas CNN como extratores de atributos, seguido por duplo estágio de seleção e classificação. O *dataset* foi uma composição de três bases de imagens obtidas via Github e Kaggle. Essas imagens são fornecidas a três CNNs – VGG19, Xception e Inception v3 – todas pré-treinadas no banco de imagens ImageNet e refinadas (*fine-tuning*). Os descritores *deep learned*, extraídos a partir das camadas globais *average pooling*, foram concatenados em um único vetor de atributos, um para cada arquitetura.

Em seguida, Bhowal, Sen e Sarkar (2021) realizaram a seleção do subconjunto de atributos relevantes, com a aplicação de um duplo estágio de seleção. A primeira etapa utilizou a teoria dos jogos “*Coalition Game theory*”, aplicando os *Shapley Values*. Basicamente, o método foi utilizado para entender a contribuição que cada atributo tem no processo de classificação. Os 1000 atributos com maiores valores de *Shapley Values* foram dados como entrada ao segundo método de seleção, *Nystrom sampling*, utilizando a decomposição CUR. Essa segunda etapa de seleção ajudou a transformar o espaço dimensional de atributos pré-selecionado em um espaço com dimensão inferior. O classificador utilizado foi o *Multilayer Perceptron*, que apresentou acurácia de 93,33%, via atributos das redes VGG-16 e Inception v3, e 94,17% com as características da Xception. Para fins de comparação com a literatura, os autores aplicaram o

método em outra base de imagens (COVIDx), obtendo acurácia de 93,45%.

O trabalho elaborado pelos autores Candelero et al. (2020) visou combinar os descritores *handcrafted* (Haralick e técnicas fractais) e *deep learned* (LeNet) para a análise de imagens histológicas de câncer colorretal, de mama e linfomas não-Hodgkin, a fim de investigar a melhor combinação entre seletor e classificador. Três arquiteturas CNN foram utilizadas para a classificação das bases de imagens: AlexNet, VGGNet e LeNet. A arquitetura que indicou o melhor desempenho, LeNet, forneceu as características *deep learned*. Os 50 melhores atributos do vetor, segundo o algoritmo ReliefF, foram dados como entrada aos algoritmos de seleção bioinspirados: GA, PSO e bGWO. Os classificadores utilizados foram SVM, *Random Forest* e K*. A partir dos experimentos, concluiu-se que a melhor combinação entre seletor e classificador foi GA+K* obtendo, além de conjuntos com menor quantidade de características, as melhores acurácias, em geral (colorretal – 91,06%; câncer de mama – 90,52%; linfomas não-Hodgkin – 82,01%). Os autores também concluíram que a combinação proposta conseguiu melhores resultados que os conquistados com as técnicas aplicadas isoladamente.

3.8 Considerações sobre os Trabalhos Relacionados

As características *handcrafted* desempenham a função de descrever quantitativamente as informações visuais que constituem as imagens digitais. Neste contexto, é possível verificar na literatura uma vasta combinação de técnicas para o reconhecimento de padrões e análise das imagens, nos mais diferentes contextos. Quando consideradas medidas de texturas, por exemplo, os descritores de Haralick, unicamente, se mostraram relevantes para a detecção do câncer de próstata e colorretal, como mostrado pelos autores Wibmer et al. (2015) e Caruso et al. (2018), respectivamente. Para potencializar o poder de descrição dessas características, a associação de diversos descritores também foram explorados. Por exemplo, Ribeiro et al. (2019) utilizaram, além

dos descritores de Haralick, múltiplas técnicas fractais sob abordagem multidimensional e multiescala para a diferenciação dos casos de câncer colorretal. Os autores Ivanovici, Richard e Decean (2009) mostraram que as técnicas de dimensão fractal e lacunaridade estendidas para o domínio das imagens coloridas apresentaram bons resultados na diferenciação de lesões psoriáticas. Da mesma forma, além da abordagem multidimensional, os autores Roberto et al. (2019) aplicaram a técnica de percolação sob ótica multiescala para a análise de tumores de mama e colorretais, indicando que atributos propostos ofereceram bons resultados na diferenciação entre os grupos das doenças analisadas.

Mesmo com os avanços indicados previamente, no trabalho de Nanni et al. (2020) foi possível constatar também que a associação entre características *handcrafted* tem melhor desempenho quando comparado a cada uma delas separadamente. Apesar disso, os autores concluíram que características *handcrafted* associadas com as obtidas pelas redes neurais convolucionais conseguiram discriminar melhor as classes sob análise. Embora os atributos extraídos diretamente das camadas mais profundas das CNNs indicarem resultados relevantes, como demonstrado nos trabalhos de Spanhol et al. (2017) e Araújo et al. (2017), o *ensemble* de características *handcrafted* e *deep learning* apresentam resultados expressivos. Por exemplo, Nanni et al. (2020) realizaram diferentes combinações de descritores e redes para descobrir a melhor combinação de técnicas capazes de classificar as imagens em estudo, assim como os autores Hassan et al. (2022).

Ainda, dada a alta dimensionalidade do espaço de atributos, impactado principalmente pelas *deep learned features*, estratégias foram desenvolvidas a partir do duplo estágio de seleção de atributos, visando reduzir a dimensionalidade e aumentar o desempenho do sistema proposto, conforme demonstrado no trabalho dos autores Bhowal, Sen e Sarkar (2021). Neste exemplo, a dupla seleção foi aplicada ao *ensemble* de descritores *deep learned*, que indicaram melhores resultados quando comparado a

um único estágio de seleção. Já os autores Candelero et al. (2020) aplicaram a técnica no *ensemble* de descritores *deep learned* e *handcrafted*, obtendo resultados relevantes na análise de imagens histológicas. A utilização dessa técnica não é restrita somente ao domínio de imagens, como apresentado no trabalho dos autores Zhang et al. (2018), que aplicam a metodologia no domínio da frequência, demonstrando a versatilidade da abordagem.

Para sumarizar o contexto analisado, nas Tabelas 3.1 e 3.2 estão alguns dos trabalhos mais recentes envolvendo o *ensemble* dos descritores, *handcrafted* e *deep learning*, e duplo estágio de seleção de atributos, respectivamente.

Tabela 3.1: Sumarização dos principais trabalhos correlatos que utilizaram *ensemble* de descritores para análises de imagens médicas.

Autores	Características <i>Handcrafted</i>	Características <i>Deep Learned</i> (CNNs)	Contexto de Aplicação	Acurácia	<i>Ensemble</i> de Classificadores?
(ALMARAZ-DAMIAN et al., 2020)	- Baseado em cor e forma que são utilizados no método ABCD (<i>Asymmetry Borders-Colors-Dermatoscopic Structures</i>) - Descritores de Haralick	VGG-16/VGG-19, MobileNet, ResNet-50, Inception v3, Xception, and DenseNet-201	- Lesão na pele (classificação binária: nevos e melanoma)	92,40%	Não
(CANDELERO et al., 2020)	Descritores de Haralick, técnicas fractais multidimensionais e multiescala	LeNet	- Câncer de Mama - Câncer Colorretal - Linfomas não-Hodgkin	- Câncer de Mama: 90,52% - Colorretal: 91,06% - Linfoma: 82,01%	Não
(HAGERTY et al., 2019)	Baseados em pigmentos atípicos, distribuição de cores, vasos sanguíneos e características demográficas	ResNet-50	- Melanoma	AUC: 0,94	Sim
(HASSAN et al., 2022)	Descritores de Haralick, histograma colorido, RSHD LDEP, SURF, DSIFT	ResNet-18, ShuffleNet, Inception-ResNet v2, Xception	Câncer de Mama	96,97%	Não
(LI et al., 2018)	RGB, HSL, SIFT, LBP, solidez, circularidade, proporção da imagem e razão da área	ResNet-50, Dense-201	- Lesão na pele (sete classes)	85,5%	Sim
(NANNI; GHIDONI; BRAHNM, 2018)	LTP, MLPQ, CLBP, RIC, FBSIF, AHP, GOLD, HOG, MOR, CLM, LET	AlexNet, GoogleNet, VGG-16, VGG-19, ResNet-50, ResNet-101, InceptionV3, IncResv2	- Linfoma - Carcinoma mamário - Laringe - Colorretal	Considerando o vetor ensemble: Carcinoma Mamário: 95,00% Linfoma: 97,33% Laringe: 95,45% Colorretal: 97,60%	Sim
(NANNI et al., 2020)	LTP, MLPQ, CLBP, RIC, GOLD, HOG, AHP, FBSIF, COL, MOR, CLM, LET	AlexNet, GoogleNet, VGG-16, VGG-19, ResNet-50, ResNet-101, Inception, IncResv2, DenseNet-201	- Células cervicais - Linfoma - Colorretal - Carcinoma mamário - Laringe	- Célula Cervical (Base 1): 94,88% - Célula Cervical (Base 2): 89,48% - Linfoma: 97,33% - Colorretal: 97,50% - Carcinoma Mamário: 96,33% - Laringe: 95,45%	Sim
(RAJINIKANTH et al., 2020)	Transformada <i>counterlet</i> , <i>curvelet</i> e transformada discreta <i>wavelet</i>	AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-101	Resonância magnética para detecção de tumor no cérebro (normal e tumor)	- Flair >99% - T2 >98% - TIC >97%	Não
(ROBERTO et al., 2021)	Percolação, lacunaridade e dimensão fractal multidimensional e multiescala	ResNet-50	- Câncer de Mama - Câncer Colorretal - Linfoma não-Hodgkin - Tecido do Fígado	- Câncer de Mama: 89,66% - Colorretal: 99,39% - Linfoma: 95,55% - Fígado: 99,62%	Sim
(SETHY; BEHERA, 2022)	GLCM, HOG, LBP	AlexNet, VGG-16, VGG-19	- Câncer de Mama (4 classes)	84,20%	Não

Portanto, neste cenário, uma investigação de *ensemble learning* para o contexto de amostras histológicas tingidas com Hematoxilina e Eosina pode contribuir significativamente para a análise, estudo e entendimento de possíveis padrões de técnicas. Para tanto, é possível analisar *ensemble learning* ainda não explorados na literatura, com múltiplos descritores, tais como *deep learned* e *handcrafted*, e múltiplos clas-

Tabela 3.2: Visão geral dos trabalhos que utilizaram duplo estágio de seleção em diferentes contextos de análise.

Autores	Método	Seletores de Atributos	Contexto de Aplicação	Acurácia
(BHOWAL; SEN; SARKAR, 2021)	Descritores <i>deep learned</i>	<i>Shapley Values</i> + (decomposição CUR + <i>Nystrom sampling</i>)	Radiografia do tórax	- COVIDx: 93,45%
(CANDELERIO et al., 2020)	<i>Ensemble</i> de atributos <i>deep learned</i> e <i>handcrafted</i>	ReliefF + <i>wrapper</i> : bGWO, GA, PSO	Imagens Histológicas	- Câncer de Mama: 90,52% - Colorretal: 91,06% - Linfoma: 82,01%
(LI et al., 2022)	Sinais de atividades elétricas cerebrais (Eletroencefalograma)	Coefficiente de correlação de Pearson + RFE (<i>recursive feature elimination</i>)	Transtorno Depressivo Maior	98,95%
(SEN et al., 2021)	Descritores <i>deep learned</i>	(Mutual Information (MI) + ReliefF) + <i>wrapper</i> : <i>Dragonfly</i> (DA)	Tomografia computadorizada	- SARS-CoV-2 CT: 98,39% - COVID-CT: 90,00%
(ZHANG et al., 2018)	Sinais de vibração	ReliefF + BPSO	Deteção de falhas para máquinas rotativas	- Falha de rolamento: 98,57% - Falha do sistema do rotor: 99,75%

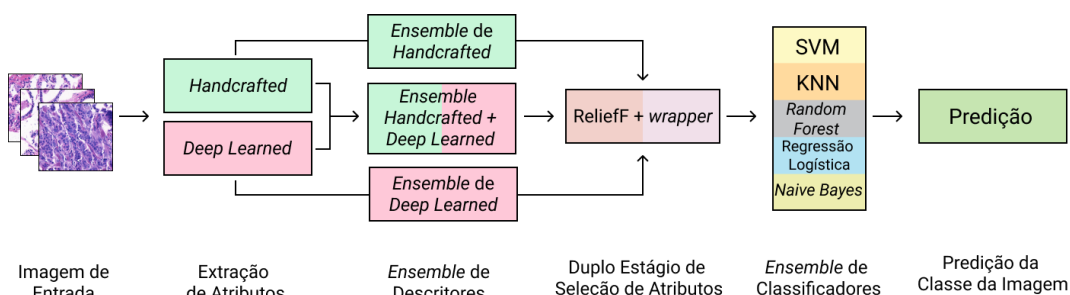
sificadores. Alguns exemplos de estruturas são: descritores *handcrafted* com um *ensemble* de classificadores; descritores *deep learned* com um *ensemble* de classificadores; descritores *handcrafted* e *deep learned* com um *ensemble* de classificadores. Este tipo de investigação possibilita responder à pertinência ou não do uso de *ensemble learning* para classificar e reconhecer padrões em alguns contextos de imagens H&E, comumente exploradas em diferentes trabalhos disponíveis na literatura.

Capítulo 4

Materiais e Métodos

O método proposto inicia com a aplicação dos descritores *handcrafted*, técnicas fractais multidimensional e multiescala, bem como Haralick e LBP. Em seguida, as características *deep learned* são obtidas utilizando diferentes arquiteturas de CNN, tais como das redes ResNet-50, VGGNet-19, Inception v3, DenseNet-121 e EfficientNet-B2. Para isso, as redes foram exploradas utilizando a estratégia de *transfer learning*. Os atributos foram utilizados para definir *ensembles* de descritores que, posteriormente, foram avaliados via algoritmos de seleção de características: ranqueamento com meta-heurísticas. Finalmente, um *ensemble* de classificadores foi definido e aplicado para testar os diferentes *ensembles* de descritores. Um esquema do método é apresentado na Figura 4.1. Todo esse processo é apresentado detalhadamente nas próximas seções.

Figura 4.1: Ilustração do método proposto.



Fonte: Elaborado pelo autor.

4.1 Descritores *handcrafted* – Técnicas fractais multidimensionais e multiescala

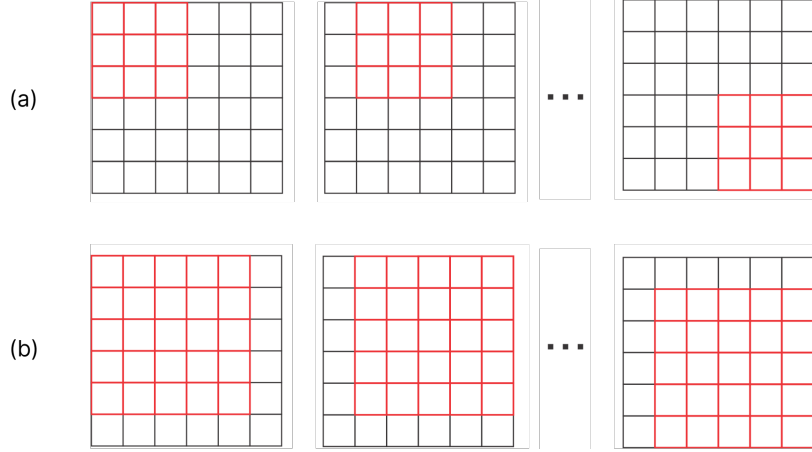
Diversas técnicas fractais podem ser encontradas na literatura para investigações de imagens. Neste trabalho, foram consideradas técnicas sob a ótica multidimensional e multiescala, tais como dimensão fractal (probabilística (IVANOVICI; RICHARD, 2011; IVANOVICI; RICHARD; DECEAN, 2009), fusão de caixas (NIKOLAIDIS; NIKOLAIDIS; TSOUROS, 2011) e Higuchi (KLONOWSKI et al., 2013; KLONOWSKI et al., 2018)), lacunaridade (IVANOVICI; RICHARD, 2011; IVANOVICI; RICHARD; DECEAN, 2009) e percolação fractal (ROBERTO et al., 2017), visto que permitem quantificações complementares de imagens coloridas.

4.1.1 Dimensão fractal baseada em abordagem probabilística

O cálculo da dimensão fractal baseada em abordagem probabilística (DF_p) foi feito como em (IVANOVICI; RICHARD, 2011; IVANOVICI; RICHARD; DECEAN, 2009). Para tanto, um *pixel* de uma imagem colorida dada como entrada, no modelo de cores *RGB*, foi representado por meio de um vetor 5D (x, y, r, g, b) , no qual as coordenadas espaciais (x, y) possuem componentes de cores (r, g, b) . Em seguida, um hipercubo de lado L foi inicialmente posicionado no canto superior esquerdo da imagem. Esse hipercubo, ao longo das iterações, foi reposicionado na imagem deslocando-se da esquerda para a direita, de cima a baixo, cobrindo todos os seus *pixels*, em um processo chamado de *gliding box*. O tamanho do hipercubo foi incrementado em duas unidades após encerrado o processo de análise, quando a estrutura atingiu o canto inferior direito da imagem. Um exemplo desse processo é ilustrado na Figura 4.2. Em sua parte superior, são ilustradas três iterações do *gliding box*, utilizando um hipercubo, representado em vermelho, tal que $L = 3$. Em seguida, essa estrutura assume valor $L = 5$ e continua o processo até varrer todos os pixels da imagem. Esse processo é repetido

até um L_{max} .

Figura 4.2: Ilustração do método *gliding box* considerando $L = 3$ (a) e $L = 5$ (b), com um hipercubo, representado em vermelho, deslocando-se na imagem até o canto inferior.



Fonte: Elaborado pelo autor.

É importante destacar que uma análise foi realizada entre os *pixels* contidos em um hipercubo. Para isso, foi fixado o *pixel* central do hipercubo, $F_c = f(x_c, y_c, r_c, g_c, b_c)$, e efetuada uma comparação com os demais, incluindo o próprio *pixel* central, baseado em uma medida de distância. Os *pixels* analisados nesse processo foram definidos como $F_i = f(x_i, y_i, r_i, g_i, b_i)$. Em nosso trabalho, consideramos análises a partir das distâncias *chessboard* (Δ_{ches}), euclidiana (Δ_{eucl}) e *manhattan* (Δ_{manh}), calculadas conforme as Equações:

$$\Delta_{ches} = \max(|F_i(o_i) - F_c(o_c)|), o \in r, g, b, \quad (4.1)$$

$$\Delta_{eucl} = \sqrt{\sum_o (F_i(o_i) - F_c(o_c))^2}, o \in r, g, b, \quad (4.2)$$

$$\Delta_{manh} = \sum_o |F_i(o_i) - F_c(o_c)|, o \in r, g, b. \quad (4.3)$$

Neste processo, cada *pixel* F_i com uma distância Δ menor ou igual ao tamanho da

escala L foi rotulado como 1 para indicar pertencente ao hipercubo. Caso contrário, o rótulo atribuído foi 0. A contagem desses *pixels* permitiu construir uma matriz de probabilidades $P(m,L)$, que caracterizou a probabilidade de m pontos estarem contidos no hipercubo de lado L . O resultado foi uma estrutura conforme a ilustração disponível na Tabela 4.1.

Tabela 4.1: Ilustração de uma matriz de probabilidades.

	3	5	...	L_{max}
1	$P(1,3)$	$P(1,5)$	...	$P(1,L_{max})$
2	$P(2,3)$	$P(2,5)$	...	$P(2,L_{max})$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
L^2	$P(L^2,3)$	$P(L^2,5)$	$\vdots$	$P(L^2,L_{max})$

A matriz $P(m,L)$ foi normalizada segundo a Equação 4.4, garantindo que a soma dos elementos de uma coluna seja igual a 1:

$$\sum_{m=1}^{L^2} P(m,L) = 1, \forall L. \quad (4.4)$$

Com a construção da matriz $P(m,L)$, foi possível obter a dimensão fractal parcial $N(L)$, que está associada a cada tamanho de hipercubo, Equação 4.5.

$$N(L) = \sum_{m=1}^{L^2} \frac{P(m,L)}{m}. \quad (4.5)$$

Calculado o valor de $N(L)$ para cada L , o coeficiente angular da regressão linear, definida por $\log L \times \log N(L)$, permitiu obter a dimensão fractal probabilística DF_p da imagem.

4.1.2 Dimensão fractal baseada em abordagem de fusão de caixas

A dimensão fractal baseada em abordagem de fusão de caixas (DF_n) foi aplicada conforme descrição apresentada por (NIKOLAIDIS; NIKOLAIDIS; TSOUROS, 2011), visando obter a dimensão fractal de um espaço n -dimensional, considerando o

espaço de cores RGB. Para isso, cada eixo da imagem foi dividido em um total de s partições, sendo possível construir uma tabela de partições. A partição em um eixo, por exemplo, o eixo de coordenadas x , se deu por meio da seguinte equação:

$$t_x = \lfloor \frac{x}{\epsilon_x} \rfloor = \lfloor \frac{xs}{L_x} \rfloor, \quad (4.6)$$

em que t_x indica a partição do eixo x , x é a coordenada de qualquer *pixel* na caixa, L_x é a escala de observação e ϵ_x refere-se à relação L/s .

A tabela obtida permitiu armazenar as coordenadas de todas as partições que contém ao menos um elemento do conjunto de dados. A cada iteração do método foi contabilizada a quantidade z de linhas da tabela, todas distintas, visto que linhas idênticas foram fundidas. Assim, a partir do coeficiente angular do gráfico de $\log z \times \log s$, foi possível definir o valor de DF_n (NIKOLAIDIS; NIKOLAIDIS; TSOUROS, 2011).

4.1.3 Dimensão fractal de Higuchi multidimensional e multiescala

A dimensão fractal proposta por Higuchi (DF_H) em 1988 (HIGUCHI, 1988) permite analisar séries temporais, domínio 1D. Dessa forma, o método mostrou-se relevante para análises de sinais biológicos (GOMES et al., 2017; GOMOLKA et al., 2018; VARLEY et al., 2020). Em nosso trabalho, o modelo utilizado foi o descrito em (TENGUAM et al., 2020), que expandiu a análise com a inclusão de estratégia multidimensional e multiescala.

Para tanto, a estratégia foi considerar séries de dados 1D de uma imagem I , com dimensões $h \times w$. A etapa multiescala, de modo geral, foi definida por meio de diferentes escalas de observações. Isso foi computado considerando vetores deslizantes que percorreram cada série de *pixels*, desde a sua primeira posição até atingir a última. A primeira posição do vetor deslizante foi definida a partir da primeira posição da série de *pixels* sob análise. Em seguida, um vetor auxiliar foi incrementado em uma posição para a direita. Esse processo foi repetido até que sua última posição coincida com a

última posição da série. Os comprimentos dos vetores foram representados por l , com valores ímpares, tal que $3 \leq l \leq \min(h,w)$, dada a necessidade de cada vetor deslizando ter um *pixel* central.

Neste processo, cada vez que o vetor se deslocou na série, foi feita uma análise entre *pixels*, caracterizando a abordagem multidimensional, estratégia que permitiu quantificar os padrões presentes em imagens coloridas. Para tanto, conforme já indicado previamente, um *pixel* de uma imagem colorida foi entendido como um vetor 5D, do tipo (x,y,r,g,b) (IVANOVICI; RICHARD; DECEAN, 2009; TENGUAM et al., 2020).

O cálculo da dimensão fractal de Higuchi (DF_H) foi definido considerando cada subsérie auxiliar finita extraída das séries de entrada da imagem I . Assim, dada uma série finita de pontos discretos: $X = x(1), x(2), x(3), \dots, x(N_{DH})$, tal que x foi um elemento obtido da imagem e N_{DH} o número máximo de pontos disponíveis na série em análise, o método construiu d novas séries X_{dj} , cujo ponto inicial se deu em j e com d sendo um fator de granularidade:

$$X_{dj} : x(j), x(j+d), x(j+2d), \dots, x\left(j + \left\lfloor \frac{N_{DH}-j}{d} \right\rfloor \cdot d\right), \quad (4.7)$$

com $j = 1, 2, \dots, d$.

Para cada X_{dj} , foi calculado o comprimento da curva, $Len_j(d)$, tal que:

$$Len_j(d) = \frac{1}{d} \left\{ \left(\sum_{i=1}^{\left\lfloor \frac{N_{DH}-j}{d} \right\rfloor} |x(j+id) - x(j+(i-1) \cdot d)| \right) \frac{N_{DH}-1}{\left\lfloor \frac{N_{DH}-j}{d} \right\rfloor \cdot d} \right\}. \quad (4.8)$$

Assim, tem-se Len como o comprimento da curva para o intervalo d , calculando o valor médio sobre os d conjuntos de $Len_j(d)$:

$$Len(d) = \frac{1}{d} \sum_{j=1}^d Len_j(d). \quad (4.9)$$

O valor DF_H para uma subsérie foi dado pelo coeficiente angular da regressão linear de uma parcela logarítmica dupla de $\ln Len(d)$ e $\ln d$, obtido por meio do ajuste linear de mínimos quadrados.

As t subséries auxiliares de *pixels* de cada conjunto de entrada inicial foram submetidas ao cálculo de DF_H (TENGUAM et al., 2020). Assim, existem t valores de DF_H . Definiu-se o valor de dimensão fractal da série de entrada, $DF_{HSérie}$, como a média desses t valores:

$$DF_{HSérie} = \frac{1}{t} \sum_{i=1}^t DF_{H_i}, \quad (4.10)$$

tal que t indica a quantidade de valores que l pode assumir.

Semelhantemente, o valor de dimensão fractal para a imagem colorida foi dada pela média dos valores de todas as $DF_{HSérie}$ (TENGUAM et al., 2020).

4.1.4 Lacunaridade

O método multidimensional e multiescala para o cálculo da lacunaridade (*Lac*) das imagens em estudo foi baseado na abordagem de (IVANOVICI; RICHARD, 2009; IVANOVICI; RICHARD; DECEAN, 2009), que utilizaram a mesma matriz de probabilidade calculada no método probabilístico da dimensão fractal, descrita na Seção 4.1.1. A métrica foi baseada em momentos de primeira ordem (Equação 4.11) e segunda ordem (Equação 4.12), segundo a medida de distribuição dada pela Equação 4.13:

$$\lambda(L) = \sum_{m=1}^{L^2} mP(m,L), \quad (4.11)$$

$$\lambda^2(L) = \sum_{m=1}^{L^2} m^2P(m,L), \quad (4.12)$$

$$\Lambda(L) = \frac{\lambda^2(L) - (\lambda(L))^2}{(\lambda(L))^2}. \quad (4.13)$$

4.1.5 Percolação multidimensional e multiescala

A percolação (*Perc*) multidimensional e multiescala foi calculada aplicando a estratégia descrita em (ROBERTO et al., 2017). Para tanto, a teoria da percolação foi explorada para a análise de caminhos de *pixels* conectados entre uma extremidade e outra da imagem. O método aplicado considerou uma abordagem multiescala com o uso da técnica *gliding box*. Para tanto, os hipercubos foram definidos inicialmente com $L = 3$ (RIBEIRO et al., 2019; ROBERTO et al., 2017), incrementado em duas unidades após percorrer a imagem. A relação da quantidade T de hipercubos que percorreram uma imagem com altura H e largura W , em função de L , foi dada por:

$$T(L) = (H - L + 1) \times (W - L + 1), L \leq \min(H, W). \quad (4.14)$$

Para cada hipercubo de tamanho L , foi aplicada uma abordagem multidimensional que considerou o canal de cor com maior intensidade, segundo o modelo RGB, visando realizar uma comparação em relação ao *pixel* central P_c , conforme apresentado em (IVANOVICI; RICHARD, 2011; IVANOVICI; RICHARD; DECEAN, 2009), Seção 4.1.1. A comparação também foi definida a partir das três distâncias mencionadas nas Equações 4.1 a 4.3. Logo, quando a distância Δ assumiu um valor menor ou igual a L , o *pixel* P foi rotulado com -1, indicando que o *pixel* representa um poro.

Os aglomerados de percolação foram obtidos considerando a rotulagem do algoritmo Hoshen-Kopelman, conforme descrição apresentada em (ROBERTO et al., 2017). Dada essa rotulagem, um aglomerado foi definido considerando *pixels* vizinhos com o mesmo rótulo. Quando o primeiro aglomerado foi identificado, o algoritmo avança para o próximo poro não verificado. A partir desse processo, três funções foram extraídas: a média de aglomerados C ; a razão de caixas percolantes Q ; e a razão

de cobertura média do maior aglomerado M . Para o cálculo da média de aglomerados por caixa $C(L)$, utilizou-se o número de aglomerados em uma única caixa (c_i), em função de uma escala L , dividido pelo total de caixas T , conforme a equação:

$$C(L) = \frac{\sum_{i=1}^{T(L)} c_i}{T(L)}. \quad (4.15)$$

O valor da razão de caixas percolantes Q foi obtida contando o número de caixas percoladas, segundo a escala L . Uma caixa q_i foi considerada percolante se a razão entre o número de *pixels* rotulados como poros (Ω_i) e o total de *pixels* na caixa (L^2) excedeu o limiar de percolação $p = 0,59275$:

$$q_i = \begin{cases} 1, & \frac{\Omega_i}{L^2} \geq p \\ 0, & \frac{\Omega_i}{L^2} < p. \end{cases} \quad (4.16)$$

A razão de caixas percolantes em função de L ($Q(L)$) foi obtida a partir da divisão do número total de caixas percolantes q_i pelo número total de caixas T em uma escala L , expressão dada por:

$$Q(L) = \frac{\sum_{i=1}^{T(L)} q_i}{T(L)}. \quad (4.17)$$

Por fim, a razão de cobertura média do maior aglomerado (M) foi calculada identificando a razão de cobertura do maior aglomerado em cada caixa avaliada na escala L , conforme a Equação 4.18, em que γ_i indicou o maior aglomerado de uma caixa i :

$$M(L) = \frac{\sum_{i=1}^{T(L)} \gamma_i / L^2}{T(L)}. \quad (4.18)$$

4.1.6 Métricas obtidas das Curvas dos descritores

Os descritores fractais baseados na abordagem probabilística, lacunaridade e percolação foram calculados com variações de escala L , segundo o método de *gliding box*.

Nesses casos, foi considerado o valor $L_{max} = 41$ como abordado em (ROBERTO et al., 2021). A dimensão fractal de Higuchi multidimensional e multiescala foi determinada com variações de escala l do vetor deslizante, também com valor máximo de 41, e um fator de granularidade $d = 8$ (TENGUAM et al., 2020). Com isso, foi possível definir curvas constituídas pelos valores assumidos em cada abordagem para cada imagem analisada. Nas abordagens de dimensão fractal de Higuchi e lacunaridade, as curvas foram formadas a partir dos seus respectivos valores locais em relação a cada tamanho de vetor/hipercubo deslizante. Já na abordagem de percolação, as curvas obtidas são C , Q e M , referente às regiões percolantes.

De maneira complementar, cada curva foi representada por valores escalares para compor os vetores de descritores (CĂLIMAN A. ; IVANOVICI, 2012; ROBERTO et al., 2017). As métricas extraídas foram: área sob a curva (A), obliquidade (O), razão de área (Γ), ponto máximo (PM) e escala do ponto (EPM). A área sob a curva (A) foi calculada pelo método de integração numérica trapezoidal indicado pela Equação 4.19.

$$A = \int_a^b f(x)dx \approx \frac{b-a}{2N} \sum_{n=a}^{b-1} (f(x_n) + f(x_{n+1})), \quad (4.19)$$

tal que a e b indicam os valores mínimo e máximo de L , respectivamente, e N o número de amostras.

A obliquidade (O) permitiu definir a assimetria de uma amostra em relação ao seu valor médio. Uma obliquidade negativa indicou que a amostra se concentrou no lado esquerdo do valor médio. Caso positivo, indicou concentração no lado direito do valor médio. Uma distribuição perfeita apontou que não houve assimetria, portanto, a obliquidade foi igual a zero. Dada uma amostra com N valores, a obliquidade foi definida pela Equação 4.20, em que $\bar{x}$ indicou o valor médio da amostra e x_i o i -ésimo valor de x :

$$O = \frac{\frac{1}{N} \sum_{i=a}^b (x_i - \bar{x})^3}{\sqrt{\left[\frac{1}{N-1} \sum_{i=a}^b (x_i - \bar{x})^2 \right]^3}}. \quad (4.20)$$

A razão da área (Γ) foi calculada como a razão entre a área da metade direita da curva em relação à área da metade esquerda da curva. Para isso, foi considerada a área sob a curva ($Area_{a,b}$) entre dois pontos a e b sobre o eixo x , como representada pela Equação 4.21:

$$Area = \int_a^b f(x) dx. \quad (4.21)$$

Dessa relação, a razão da área Γ foi definida conforme a Equação 4.22:

$$\Gamma = \frac{A_{(b/2+1,b)}}{A_{(a,b/2)}} \quad (4.22)$$

Por fim, o ponto máximo de cada função foi considerado para definir outras duas medidas, a escala de observação (EPM) e o próprio valor da função no ponto de referência (PM).

4.2 Descritores *handcrafted* – Haralick

Os descritores de Haralick foram obtidos a partir da matriz de co-ocorrência (HARALICK; SHANMUGAM; DINSTEN, 1973). A matriz de co-ocorrência permitiu verificar a transição entre dois *pixels* em níveis de cinza, u e v , considerando uma medida de distância δ e um ângulo arbitrário θ . No trabalho proposto, foram considerados os valores de parâmetros de distância $\delta = 1$ e de direção $\theta = 0^\circ, \theta = 45^\circ, \theta = 90^\circ$ e $\theta = 135^\circ$ (CANDELERO et al., 2020; RIBEIRO et al., 2019). Com os valores das matrizes de co-ocorrência, os descritores obtidos foram: (1) Segundo momento angular; (2) Contraste; (3) Correlação; (4) Soma dos quadrados; (5) Momento inverso das diferenças; (6) Somatório da média; (7) Somatório da variância; (8) Somatório da

entropia; (9) Entropia; (10) Diferença da variância; (11) Diferença da entropia; (12) e (13) são ambas medidas de informação de correlação; (14) Coeficiente de correlação maximal. Para os 14 descritores finais são consideradas as médias dos descritores obtidos nas quatro matrizes (CANDELERO et al., 2020; RIBEIRO et al., 2019).

4.3 Descritores *handcrafted* – *Local Binary Pattern* (LBP)

O método LBP foi aplicado conforme descrições apresentadas em (MAZO; ALLEGRE; TRUJILLO, 2017; YU et al., 2019), considerando um raio de valor $R = 1$ e o número de vizinhos $PN = 8$. A partir desses parâmetros, foi possível obter um total de 59 atributos para cada uma das imagens em análise, segundo o método *extractLBPFatures(I)*, disponível na ferramenta Matlab (MATHWORKS, c2022).

4.4 Descritores *deep learned*

Os descritores *deep learned* foram calculados a partir do uso de transferência de aprendizado (PAN; YANG, 2010), pois permitiu reduzir o tempo de treinamento do modelo CNN e possibilitou análises envolvendo conjuntos de imagens com um número reduzido de amostras. Portanto, em nossa proposta, utilizamos a rede pré-treinada no banco de imagens ImageNet (ALMARAZ-DAMIAN et al., 2020; ROBERTO et al., 2021). Assim, conseguimos aplicar os modelos CNN exclusivamente para a extrair os valores de camadas específicas e utilizá-los como *deep learned*.

Consideramos cinco arquiteturas distintas de CNN: ResNet-50 (HE et al., 2016), VGG-19 (SIMONYAN; ZISSERMAN, 2014), Inception v3 (SZEGEDY et al., 2015b), DenseNet-121 (HUANG et al., 2017) e EfficientNet-B2 (TAN; LE, 2019). Informações adicionais de cada uma delas estão sumarizadas na Tabela 4.2. Essas redes apresentaram resultados relevantes em problemas de classificação de imagens médicas em contextos variados (ALMARAZ-DAMIAN et al., 2020; NANNI; GHIDONI; BRAH-

NAM, 2018; RAJINIKANTH et al., 2020) e também na classificação de imagens histológicas (NANNI et al., 2020; RAKHLIN et al., 2018; ROBERTO et al., 2021). Para a arquitetura ResNet-50, os descritores *deep learned* foram obtidos da última camada convolucional, no ponto imediatamente anterior à camada totalmente conectada, como explorado nos trabalhos dos autores (ALMARAZ-DAMIAN et al., 2020; HAGERTY et al., 2019; RAKHLIN et al., 2018). Com relação à arquitetura VGG-19, os atributos foram os resultantes da segunda camada totalmente conectada (KWASIGROCH; MIKOŁAJCZYK; GROCHOWSKI, 2017; SANTOS; PONTI, 2019). Nas arquiteturas Inception v3 e DenseNet-121, os atributos foram extraídos de maneira similar à ResNet-50 (camada avgpool), no ponto imediatamente anterior à última camada totalmente conectada, como nos trabalhos (BEN HAMIDA et al., 2021; KALRA et al., 2020). Por fim, os atributos *deep learned* explorados na arquitetura EfficientNet-B2 se deram a partir da última camada totalmente conectada (MUNIEN; VIRIRI, 2021).

Tabela 4.2: Resumo de cada modelo CNN investigado neste projeto.

Modelo CNN	Total de Camadas	Total de Parâmetros*	Dimensões de Entrada
ResNet-50	50	25×10^6	$224 \times 224 \times 3$
VGG-19	19	143×10^6	$224 \times 224 \times 3$
Inception v3	48	23×10^6	$299 \times 299 \times 3$
DenseNet-121	121	8×10^6	$224 \times 224 \times 3$
EfficientNet-B2	342	$9,2 \times 10^6$	$260 \times 260 \times 3$

* Fonte: (WOLFRAM, c2022)

4.5 *Ensemble de descritores*

Neste trabalho, exploramos *ensembles* de descritores *handcrafted* e *deep learned* que ampliam as estratégias já exploradas em (ALMARAZ-DAMIAN et al., 2020; KUMAR et al., 2022; NANNI; GHIDONI; BRAHNAM, 2018; PONTI JR., 2011; RAJINIKANTH et al., 2020; SETHY; BEHERA, 2022). Isso permitiu conhecer possíveis padrões de combinações de técnicas em diferentes tipos de amostras histológicas

H&E.

As estruturas foram: (i) *ensemble* de descritores *handcrafted* com um *ensemble* de classificadores; (ii) *ensemble* de descritores *deep learned* com um *ensemble* de classificadores; (iii) *ensemble* de descritores *handcrafted* e *deep learned* com um *ensemble* de classificadores. Além disso, coletamos os desempenhos de classificações realizadas somente com o uso dos modelos CNN com o propósito de verificar a pertinência ou não do uso de *ensemble learning* para classificar e reconhecer padrões em imagens representativas do câncer colorretal, displasia epitelial oral, linfomas não-Hodgkin e tecidos hepáticos.

As definições dos totais de descritores que constituem os *ensembles* foram dependentes de cada uma das categorias aqui exploradas. Os descritores *handcrafted* totalizam 462 valores, mais especificamente: 389 atributos fractais (Percolação – 225; Lacunaridade – 75; Dimensão fractal probabilística – 63; Dimensão fractal de Higuchi – 25; Dimensão fractal por fusão de caixas – 1); 59 atributos LBP e 14 descritores de Haralick. Informações mais detalhadas estão sumarizadas na Tabela 4.3. Quanto aos descritores *deep learned*, totalizam 10.624 valores. Esse total depende das camadas utilizadas nas análises. Para este trabalho, conforme descrições apresentadas na subseção 4.4, as camadas escolhidas e os valores correspondentes estão sumarizados na Tabela 4.4.

4.6 Análises de descritores via duplo estágio de seleção e *ensemble* de classificadores

Os *ensembles* de descritores foram avaliados aplicando um processo com duplo estágio de seleção de atributos para diminuir a dimensionalidade do espaço de características e identificar a melhor combinação dentre os descritores disponíveis em cada subconjunto (CANDELERO et al., 2020; HSU; HSIEH; LU, 2011; MENGDI et al.,

Tabela 4.3: Nome, medida de distância considerada e o total de características *hand-crafted* resultante.

Atributo	Distância	Número de Atributos
Percolação - C, Q, M	<i>Chessboard</i>	60
Lacunaridade		20
Dimensão Fractal Probabilística Local		20
Percolação - C, Q, M	Euclidiana	60
Lacunaridade		20
Dimensão Fractal Probabilística Local		20
Percolação - C, Q, M	<i>Manhattan</i>	60
Lacunaridade		20
Dimensão Fractal Probabilística Local		20
Métricas Percolação - C, Q, M	<i>Chessboard</i> Euclidiana <i>Manhattan</i>	45
Métricas Lacunaridade	<i>Chessboard</i> Euclidiana <i>Manhattan</i>	15
Dimensão Fractal Probabilística Global	<i>Chessboard</i> Euclidiana <i>Manhattan</i>	3
LBP	-	59
Haralick	-	14
Dimensão Fractal de Higuchi	<i>Chessboard</i>	20
Métricas Dimensão Fractal de Higuchi		5
Dimensão Fractal Fusão de Caixas	-	1
Total de Atributos		462

Tabela 4.4: Número de descritores extraídos e sua referente camada em cada arquitetura.

Arquitetura CNN	Número de descritores	Camada extraída
ResNet-50	2048	avgpool
VGGNet-19	4096	FC2
Inception v3	2048	última avgpool
DenseNet-121	1024	global avgpool
EfficientNet-B2	1408	FC

2018). O primeiro estágio da nossa proposta consistiu em ranquear cada entrada por meio do algoritmo ReliefF (URBANOWICZ et al., 2018) e aplicar um limiar para reduzir o número de possíveis combinações. O ponto de corte considerou os 100 atributos mais bem ranqueados em cada composição *ensemble* (RIBEIRO et al., 2019). Com isso, o segundo estágio foi direcionado para encontrar as melhores combinações existentes em cada subconjunto reduzido, via seleção *wrapper* e explorando diferentes meta-heurísticas, a saber: algoritmo genético (GA), baseado em evolução genética; *particle swarm optimization* (PSO), baseado em enxame de partículas; e *binary*

grey wolf optimization (bGWO), baseado em comportamento animal, mais especificamente, na estratégia de caça de lobos cinzentos (EMARYA; ZAWBAAB; HASSANIENA, 2016). Essas estratégias foram exploradas aplicando o algoritmo *K-Nearest Neighbors* para avaliar o poder as soluções de cada método de seleção. Assim, essa proposta permitiu investigar possíveis padrões de combinações de técnicas em diferentes tipos de amostras (BHOWAL; SEN; SARKAR, 2021; CANDELERO et al., 2020; LI et al., 2022). Pela característica estocástica dos algoritmos de seleção *wrapper*, o desempenho final de uma combinação foi determinado a partir da média aritmética de 10 diferentes execuções da etapa de seleção. Também, os atributos mais selecionados foram definidos por meio de suas frequências nas dez execuções (TAINO et al., 2021).

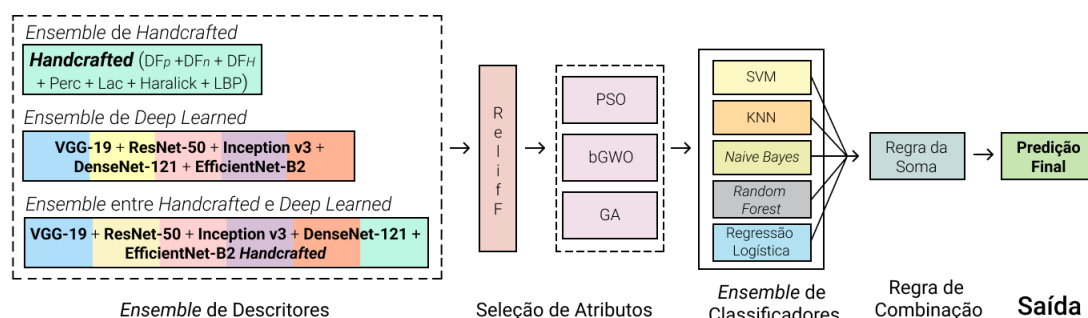
4.6.1 *Ensemble Learning*

A última etapa do processo foi direcionada para avaliar cada *ensemble* obtido na etapa anterior. Essa avaliação foi por meio de um *ensemble* de classificadores. A abordagem permitiu combinar resultados de classificadores individuais para evitar *overfitting* (HAGERTY et al., 2019; PONTI JR., 2011; ROBERTO et al., 2021). Aqui, o *ensemble* de classificadores foi constituído por algoritmos representativos de diferentes categorias: *Support Vector Machine* (SVM); KNN(K-Nearest Neighbors); *Random Forest*; Regressão Logística; *Naive Bayes* (PONTI JR., 2011).

Os classificadores foram treinados separadamente e o *ensemble* foi aplicado considerando a regra da soma (NANNI; GHIDONI; BRAHNAM, 2018; PONTI JR., 2011; ROBERTO et al., 2021). Na Figura 4.3 é apresentada uma visão ilustrativa do processo de avaliação e esquema de *ensemble*. Os blocos indicados por linhas tracejadas indicam que apenas um dos blocos internos foi utilizado na composição do modelo.

O processo de avaliação com cada algoritmo de classificação foi a partir do método *k-fold cross-validation*, com $k = 3$ (CANDELERO et al., 2020). Neste método, as instâncias são divididas em k grupos independentes, cada um particionado em um

Figura 4.3: Esquema do *ensemble* proposto em que as caixas com linhas tracejadas definem que apenas um dos seus componentes foi utilizado para compor o *ensemble*.



Fonte: Elaborado pelo autor.

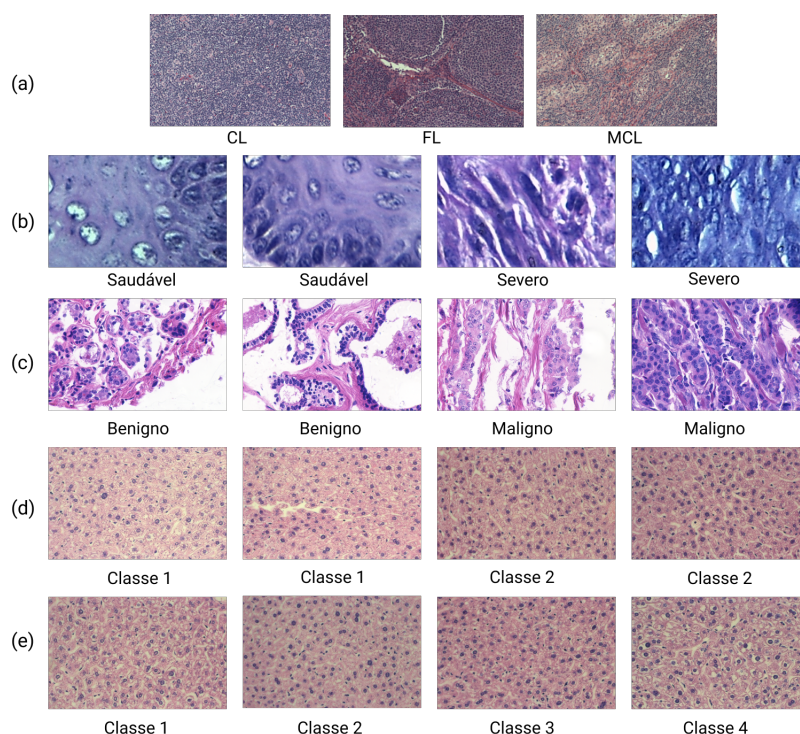
conjunto de treinamento e um conjunto de teste. As instâncias de cada grupo foram classificadas obtendo-se ao final k medidas de desempenho, uma para cada *fold*. O desempenho final foi definido pela média aritmética dos desempenhos parciais. As métricas utilizadas para avaliar cada solução são: área sob a curva ROC (AUC) e acurácia (ACC) (BRADLEY, 1997).

4.7 Contextos de aplicação

O trabalho proposto foi aplicado na avaliação de cinco bases de imagens histológicas: a) câncer colorretal (CR) (SIRINUKUNWATTANA et al., 2017); b) displasia epitelial oral (OED) (SILVA et al., 2022); c) linfomas não-Hodgkin (NHL) (SHAMIR et al., 2008); tecidos hepáticos, com d) *Liver Gender* (LG) e e) *Liver Aging* (LA) (AGEMAP, 2020). Exemplos de cada base de imagens e suas respectivas classes estão apresentadas na Figura 4.4.

O conjunto (CR) é composto por imagens histológicas derivadas de 16 seções coradas em H&E (*Hematoxylin & Eosin*) do estágio T3 ou T4 de câncer colorretal. As imagens foram rotuladas por um especialista na área de patologia, categorizadas no grupo benigno ou maligno. A base é composta por 165 imagens, divididas em 74 imagens de casos benignos e 91 imagens de casos malignos (SIRINUKUNWATTANA et

Figura 4.4: Exemplos de imagens NHL (a), OED (b), CR (c), LG (d) e LA (e).



Fonte: Elaborado pelo autor.

al., 2017).

A base (OED) foi construída a partir de 30 seções histológicas H&E da língua de camundongos previamente submetidas a um carcinógeno (SILVA et al., 2022). Esse estudo foi aprovado pelo comitê de Ética, protocolo 038/39 da Universidade Federal de Uberlândia. Um total de 148 ROIs, divididos em duas classes, saudável (74) e displasia severa (74), foram consideradas neste trabalho.

O conjunto de imagens (NHL) é constituído por 374 imagens H&E de linfomas malignos Não-Hodgkin subdivididas em três classes: Leucemia Linfóide Crônica (CLL) – 113 imagens; Linfoma Folicular (FL) – 139 imagens; e Linfoma de Células Manto (MCL) – 122 imagens. O conjunto de dados é uma coleção de amostras preparadas por diferentes histologistas em vários hospitais (SHAMIR et al., 2008).

As duas últimas bases são compostas por tecidos hepáticos de 48 ratos machos e fêmeas obtidos por meio do estudo *Atlas of Gene Expression in Mouse Aging Project*

(AGEMAP), pelo *National Institute on Aging*. As imagens foram adquiridas manualmente por um microscópio *Carl Zeiss Axiovert 200* com magnificação 40x. A base *Liver Gender* (LG) é composta por 265 imagens subdivididas nas classes masculinas (150 amostras) e feminina (115) de ratos em dieta *ad-libitum*. Já o conjunto *Liver Aging* (LA) contém 528 imagens de ratos fêmeas em quatro idades: 1 (100), 6 (115), 16 (162) e 24 (151) meses, também em dieta *ad-libitum* (AGEMAP, 2020).

Uma visão geral das bases de imagens é apresentada na Tabela 4.5, com as indicações dos nomes das bases e seus respectivos tipos, número de imagens e quantidades de classes.

Tabela 4.5: Informações gerais das bases de imagens utilizadas no trabalho.

Base de Imagens	Imagem	Número de Imagens	Número de Classes	Resolução
CR	Tumor Colorretal	165	2	567 × 430 a 775 × 522
OED	Displasia Epitelial Oral	148	2	450 × 250
LA	Tecido Hepático	528	4	417 × 312
LG	Tecido Hepático	265	2	417 × 312
NHL	Linfoma Não-Hodgkin	374	3	1388 × 1040

4.8 Pacotes de *software* e ambiente para os experimentos

Neste trabalho, os descritores *handcrafted* foram implementados e extraídos via pacote MATLAB[®] R2019a (MATHWORKS, 2021). As arquiteturas de redes neurais convolucionais pré-treinadas foram implementadas por meio da biblioteca Pytorch (PYTORCH, 2021) no ambiente de execução Google Colaboratory (GOOGLE, 2021). Os algoritmos de seleção e classificação de atributos foram empregados pelo *software* Weka 3.8 (HALL et al., 2009). Foram utilizados os valores padrões sugeridos em cada ferramenta, exceto aqueles específicos mencionados durante o texto. Os experimentos foram realizados em um notebook Intel Core i5, 8 GB de memória RAM e sistema operacional com arquitetura de 64 *bit*.

4.9 Considerações Finais sobre Materiais e Métodos

Neste capítulo foi apresentada a metodologia desenvolvida para a execução do projeto. Os métodos de dimensão fractal selecionados também foram abordados em trabalhos de outros autores para a análise de imagens histológicas nos contextos de imagens aqui explorados, tais como linfomas, câncer colorretal, tecidos hepáticos e mamários. Os atributos de textura LBP e Haralick são as abordagens mais clássicas e amplamente difundidas na literatura. As arquiteturas CNN representam a categoria capaz de extrair padrões das imagens de entrada, a partir de um processo de treinamento. Esses descritores foram utilizados para a composição de diferentes combinações de *ensembles*, os quais são posteriormente analisados via duplo estágio de seleção e classificados por meio de um *ensemble* entre cinco classificadores de diferentes categorias.

Capítulo 5

Resultados e Discussão

O modelo proposto foi testado em cinco conjuntos de imagens histológicas H&E, conforme descrição apresentada no Capítulo 4. As comparações foram realizadas envolvendo as diferentes classes de cada conjunto. É importante destacar que foram realizados 45 tipos de testes para explorar distintas composições de *ensembles*, inclusive com três associações de métodos de seleção *wrapper*: GA, PSO e bGWO. Esses algoritmos buscaram as melhores composições dentre os 100 descritores mais bem ranqueados via ReliefF, conforme metodologia proposta. As composições foram avaliadas via um *ensemble* heterogêneo de classificadores (SVM, KNN, *Random Forest*, Regressão Logística e *Naive Bayes*), conforme detalhado na Seção 4.6.1 do Capítulo 4. Na Tabela 5.1 estão as médias dos desempenhos para cada *ensemble*, considerando os descritores *deep learned features* (DL) e *handcrafted features* (HF). As melhores taxas estão destacadas em negrito. As acurácias e outras informações adicionais, obtidas em cada experimento, estão disponíveis nas Tabelas A.1 a A.5 do Apêndice A.

A partir da Tabela 5.1, é possível verificar que os *ensembles* HF e HF+DL foram responsáveis pelos melhores resultados em quatro conjuntos H&E (OED, LA, LG e NHL) dos cinco investigados neste trabalho. As acurácias foram de 90,72% a 100%. Logo, é possível indicar que o *ensemble* HF, aqui definido, é relevante para o processo de classificação em diferentes tipos de imagens H&E, seja utilizado separadamente ou

Tabela 5.1: Acurácias médias obtidas pelo método proposto, utilizando dois estágios de seleção (ReliefF + *Wrapper*). As melhores acurácias estão destacadas em negrito.

<i>Ensemble</i>	<i>Wrapper</i>	CR	OED	LA	LG	NHL
HF	PSO	91,94%	100%	98,30%	99,36%	78,02%
	GA	91,39%	99,66%	98,58%	99,36%	77,59%
	bGWO	92,30%	99,66%	98,73%	99,36%	77,94%
DL	PSO	99,45%	96,89%	93,50%	98,75%	90,64%
	GA	99,15%	97,64%	93,01%	98,75%	90,56%
	bGWO	99,76%	96,96%	93,39%	98,87%	90,13%
HF+DL	PSO	99,58%	98,45%	96,86%	99,40%	90,51%
	GA	99,45%	98,45%	96,72%	99,36%	90,72%
	bGWO	99,45%	99,05%	96,44%	99,47%	90,43%

por meio de uma combinação com DL. Também, o *ensemble* HF proporcionou a maior taxa de distinção na base OED, independente do seletor *wrapper* tomado como referência, indicando a combinação ideal nesta base, o que pode representar mais uma contribuição do nosso estudo. Por outro lado, HF apresentou as menores acurácias (aproximadamente 78%) no conjunto NHL, com três classes (CLLxFLxMCL), mostrando que esta combinação não foi a mais apropriada para o reconhecimento de padrões de NHL. Entretanto, quando esta categoria foi combinada com DL, *ensemble* HF+DL, o resultado conquistado foi o mais expressivo para o conjunto NHL, com uma acurácia de 90,72%. Apesar disso, este resultado representa uma acurácia de, no mínimo, 8% menor em relação aos valores conquistados em outras bases. Este é um indicativo importante sobre as dificuldades existentes para distinguir CLLxFLxMCL, em especial com o modelo HF. Por fim, quando o *ensemble* DL é considerado, a melhor solução foi conquistada em um único conjunto (CR), mas com uma taxa expressiva (99,76% de acurácia), ilustrando sua importância para o desenvolvimento de estratégias de apoio ao diagnóstico do câncer colorretal. Adicionalmente, para este conjunto, é interessante destacar mais uma solução, o *ensemble* HF+DL, que conseguiu conquistar uma acurácia (99,58%), muito próxima da fornecida pela combinação DL, configurando-se com soluções aceitáveis e comuns para diferentes tipos de amostras histológicas.

Visando sumarizar os resultados aqui discutidos, as melhores combinações de descritores e algoritmos de seleção são apresentadas na Tabela 5.2, incluindo o total de descritores, médias de AUC e acurácia. As acurácias das soluções *top 1* e *top 10* também são apresentadas, possibilitando observar a variação existente para a primeira e décima soluções em cada base, visto que as médias foram calculadas a partir das 10 composições mais bem ranqueadas em cada conjunto H&E. É importante destacar que o ranqueamento considerou as maiores acurácias com os menores números de descritores.

Tabela 5.2: Principais combinações em cada conjunto H&E a partir do critério maior acurácia com o menor número de descritores.

Base	Ensemble	Seleção	[Média de Descritores]	Acurácia Média (%)	AUC Média	Acurácia <i>top 1</i>	Acurácia <i>top 10</i>
CR	DL	ReliefF + bGWO	21 ± 3	99,76% ± 0,30%	1	100%	99,39%
OED	HF	ReliefF + PSO	11 ± 3	100%	1	100%	100%
LA	HF	ReliefF + bGWO	40 ± 7	98,73% ± 0,36%	0,999 ± 0,4 × 10 ⁻³	99,24%	98,11%
LG	HF+DL	ReliefF + bGWO	29 ± 6	99,47% ± 0,30%	1	100%	98,87%
NHL	HF+DL	ReliefF + GA	53 ± 5	90,72% ± 1,04%	0,980 ± 1,5 × 10 ⁻³	92,25%	89,57%

A partir do critério exposto previamente, é possível reiterar que os *ensembles* HF e HF+DL foram os responsáveis pelos melhores resultados em quatro dos cinco conjuntos H&E aqui investigados. Nestes casos, destacam-se os processos de seleção ReliefF + bGWO, com três ocorrências. Este fato pode indicar outro padrão para os conjuntos CR, LA e LG, com taxas médias expressivas de AUC, 0,999 (LA) e 1 (CR e LG). Nestes conjuntos as *top 1* soluções forneceram acurácias ideais. Adicionalmente, a menor acurácia *top 10* foi expressiva, uma taxa acima de 98% (LA), base que envolve quatro grupos para o processo de classificação. Mais ainda, as principais soluções para estes três conjuntos envolveram um número reduzido de descritores, média de 21 descritores para CR, 29 para LG e 40 para LA.

Em relação ao duplo estágio ReliefF + PSO, esta estratégia constituiu a principal solução para a base OED, proporcionando acurácia máxima com o menor número de descritores dentre todas as soluções, apenas 11 atributos. Quando o duplo estágio ReliefF + GA é observado, esta abordagem constituiu a melhor solução em uma única base (NHL). Neste caso, as soluções exploraram, em média, 53 atributos, o maior va-

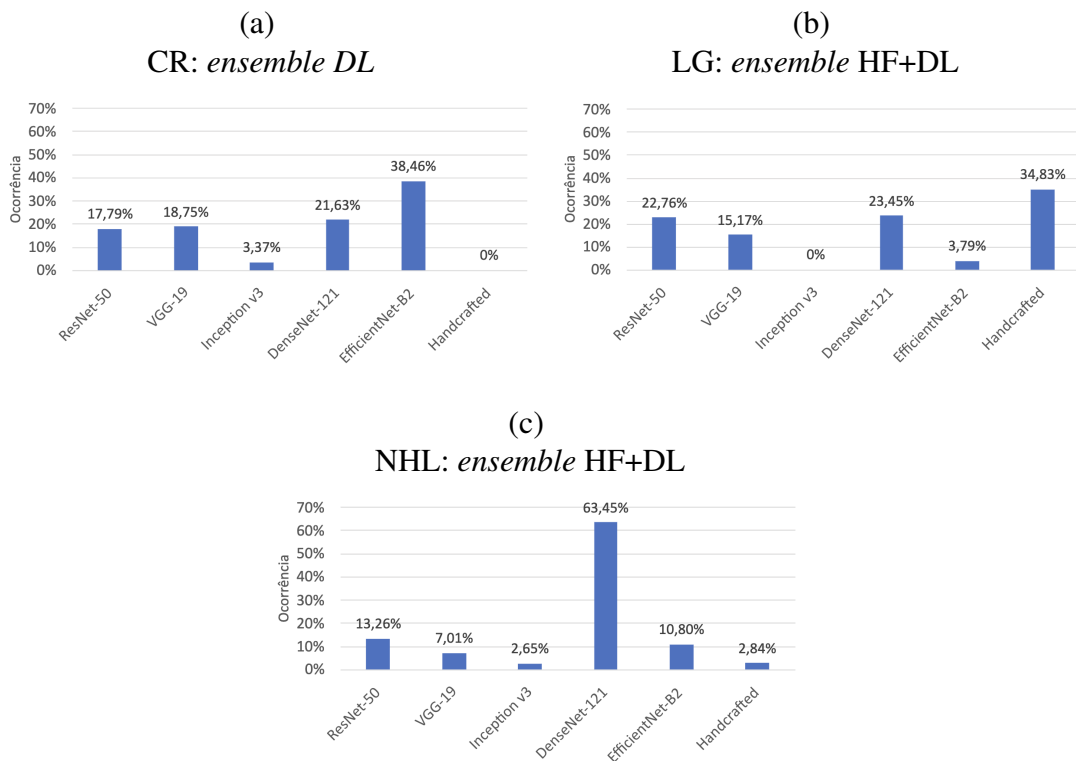
lor dentre todas as soluções. A acurácia *top* 10 foi de 89,57% e a *top* 1 foi sutilmente melhor, 92,25%, reforçando as dificuldades presentes neste conjunto. A base NHL é constituída por três classes (CLLxFLxMCL) e, possivelmente, com padrões histológicos menos heterogêneos, implicando em mais dificuldades nas constituições das soluções. Mesmo assim, neste caso, a AUC média foi de 0,98, um valor importante diante das condições expostas. Por exemplo, dado que o espaço inicial de características variou de 462 a 11.086 dimensões, a solução aqui obtida é mais uma contribuição deste estudo, capaz de proporcionar desempenhos médios expressivos com poucos descritores, mas altamente relevantes para o processo de classificação.

5.1 Ocorrências dos descritores nas principais soluções: uma visão geral

Para conhecer os descritores presentes nas principais soluções, conforme sumariado na Tabela 5.2, foi realizado um levantamento da ocorrência de cada categoria de descritores nas *top* 10 soluções de cada conjunto H&E. Para melhor compreensão da origem de cada descritor em cada solução, as ocorrências dos descritores *deep learned* estão na Figura 5.1, conjuntos CR, LG e NHL, e os *handcrafted* estão na Figura 5.2, soluções para LA, LG, OED e NHL. É importante observar que as ocorrências em NHL e LG também envolvem descritores *handcrafted*, em razão do *ensemble* HF+DL, justificando a reapresentação dessas bases na Figura 5.2. Ainda nestes dois conjuntos, as porcentagens de ocorrências foram calculadas em relação aos totais de atributos HF no *ensemble* HF+DL, desconsiderando as porcentagens das *deep features*.

Considerando as distribuições ilustradas na Figura 5.1, é possível verificar alguns comportamentos. As menores ocorrências ocorreram em relação aos descritores da rede Inception v3, máximo de 3,37% para o conjunto CR e nenhuma ocorrência dentre as *top* 10 soluções para a base LG. Por outro lado, os descritores das redes DenseNet-

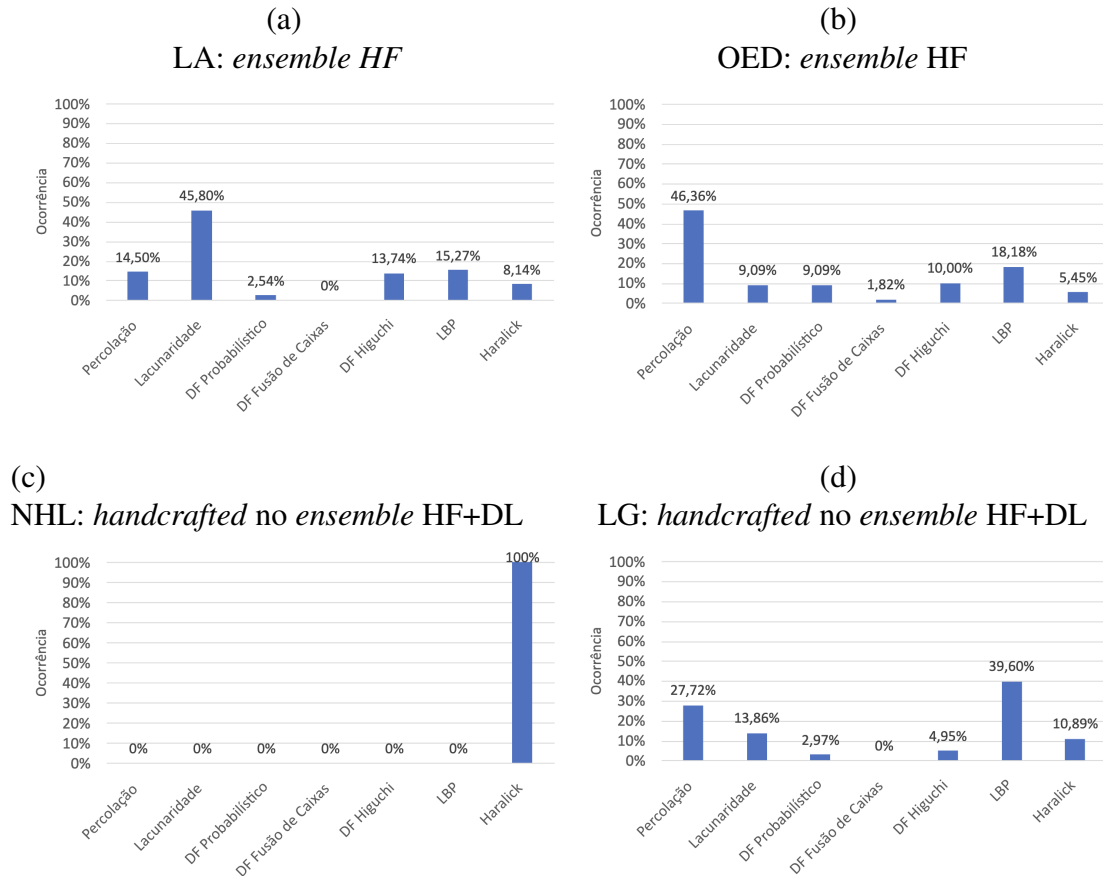
Figura 5.1: Ocorrências de descritores *deep learned* presentes nas *top 10* soluções das bases CR, LG e NHL.



121 e EfficientNet-B2 têm as maiores ocorrências, com destaque para o conjunto NHL, em que 63,45% dos atributos foram da DenseNet-121. Os descritores da EfficientNet-B2 descaram-se nas soluções para o conjunto CR, com 38,46% das ocorrências, diante de ocorrências mais homogêneas (de 17,79% a 21,63%) para os descritores das redes ResNet-50, VGG-19 e DenseNet-121. Outra distribuição homogênea pode ser constatada no conjunto LG, envolvendo as mesmas origens de descritores para a base CR. Neste caso, as ocorrências foram de 15,17% a 23,45%. Quando a totalização DL contra *handcrafted* é considerada, nota-se que descritores DL predominaram nas soluções para os conjuntos LG e NHL, com ocorrências de 65,17% e 97,16%, respectivamente. Apesar destas diferenças, não é possível indicar que estes foram os mais importantes nas classificações.

Em relação às ocorrências de descritores *handcrafted*, Figura 5.2, a menor ocorrência foi do descritor dimensão fractal por fusão de caixas (DF_n de (NIKOLAIDIS;

Figura 5.2: Ocorrência dos descritores *handcrafted* nas *top 10* soluções para as bases LA, OED, NHL e LG.



NIKOLAIDIS; TSOUROS, 2011)), visto que não foi selecionado para as *top 10* soluções em três das quatro bases H&E. Este descritor esteve presente nas soluções para OED, mas com a menor ocorrência, apenas 1,84%. O descritor dimensão fractal probabilístico (DF_p de (IVANOVICI; RICHARD, 2011; IVANOVICI; RICHARD; DECEAN, 2009)) foi o segundo com a menor ocorrência, mas constituiu as soluções para três das quatro bases histológicas. Outro comportamento interessante envolve a versão aprimorada do descritor dimensão fractal de Higuchi (DF_H (TENGUAM et al., 2020)), com ocorrências que superaram as das abordagens DF_n e DF_p (NIKOLAIDIS; NIKOLAIDIS; TSOUROS, 2011; IVANOVICI; RICHARD, 2011; IVANOVICI; RICHARD; DECEAN, 2009), amplamente exploradas na literatura. Esta é outra importante contribuição deste estudo. Por fim, é possível observar os descritores com as

maiores ocorrências para cada conjunto H&E: lacunaridade (45,80%) para LA; percolação (46,36%) para OED; LBP (39,60%) para LG; e, com destaque, Haralick como os únicos que constituíram as soluções para NHL.

5.2 Visão Geral frente a diferentes abordagens

Os melhores resultados foram observados em relação às taxas de distinções das redes (ResNet-50 (HE et al., 2016), VGG-19 (SIMONYAN; ZISSERMAN, 2014), Inception v3 (SZEGEDY et al., 2015b), DenseNet-121 (HUANG et al., 2017) e EfficientNet-B2 (TAN; LE, 2019)) aplicadas diretamente nas imagens histológicas H&E. Para tanto, utilizamos: um processo de refinamento (*fine-tuning*); validação cruzada *k-fold cross validation*, com $k = 3$; 10 épocas; algoritmo gradiente estocástico descendente (*sgdm*); *learning rate* inicial de 0,01, com decaimento de 0,75 a cada duas épocas; e, função de *loss* como *cross-entropy*. Experimentos similares foram descritos por (ROBERTO et al., 2021). Além disso, as imagens de entrada foram normalizadas segundo os valores de média e desvio padrão da base ImageNet (PYTORCH, 2022). As acurácias fornecidas pelas redes estão na Tabela 5.3.

Tabela 5.3: Acurácias fornecidas pelas arquiteturas CNN após processo de *fine-tuning*: os melhores resultados estão em negrito.

Arquitetura CNN	CR	OED	LA	LG	NHL
ResNet-50	96,73%	96,00%	91,48%	97,78%	74,27%
VGG-19	98,67%	92,67%	77,96%	91,98%	65,07%
Inception v3	85,64%	90,50%	85,37%	92,59%	67,07%
DenseNet-121	97,45%	93,17%	87,41%	93,98%	72,40%
EfficientNet-B2	96,12%	94,67%	90,00%	98,89%	70,27%

Para observar as diferenças entre as taxas de distinções dos modelos aqui propostos em relação às redes aplicadas diretamente, consideramos as acurácias médias conquistadas em cada conjunto H&E, conforme resumo indicado na Tabela 5.2. Logo, é possível verificar que as acurácias via os modelos propostos superaram as fornecidas

pelos redes ResNet-50, VGG-19, Inception v3, DenseNet-121 e EfficientNet-B2. As taxas de classificações com as redes convolucionais foram de 74,27% a 98,89%. Com isso, os ganhos nas acurácias variaram de 0,6% a 16%, aproximadamente. O menor ganho ocorreu no conjunto CR (0,58%) e o maior (16,45%) no NHL. Por exemplo, conseguimos elevar a taxa de classificação no *dataset* NHL, que envolve três classes, de 74,27% para 90,72%, ilustrando mais uma contribuição deste estudo.

Em relação às diferenças observadas, o teste de Friedman (DEMŠAR, 2006) foi aplicado para verificar se estas são estatisticamente relevantes. O teste de Friedman é um método estatístico não-paramétrico capaz de ranquear as soluções sob investigação, em que a melhor opção é definida na 1ª posição (DEMŠAR, 2006). Em caso de empate, o ranqueamento médio é atribuído entre as soluções. Este tipo de teste possibilita observar a variância de medidas repetidas, além de analisar se as diferenças existentes são estatisticamente significantes, via os *p-values*. Quanto menor o *p-value*, mais evidência se tem de que a diferença é estatisticamente relevante. É possível indicar que existe alguma diferença relevante quando o *p-value* é menor $< 0,05$. Em nossos testes, o *p-value* resultante foi de 0,0004, indicando que as diferenças entre as soluções são estatisticamente significantes.

Adicionalmente, o teste de Friedman ranqueou as soluções na forma de uma tabela. O resultado envolvendo os experimentos está na Tabela 5.4, com a indicação do *Friedman's Score*. Nota-se que as soluções obtidas neste estudo são as mais relevantes para cada conjunto de imagens. É importante observar que cada *dataset* representou uma diferente amostra (linha) e cada método uma coluna. O resultado foi uma ordem das melhores soluções para cada conjunto. No caso em que há empate, ranques médios são atribuídos para as soluções. Como resultado, cada linha da tabela possui uma posição no ranque, de acordo com seu desempenho em cada método. Ao final, o *Friedman's Score* foi definido como a média aritmética dos ranques das amostras de cada coluna da tabela, indicando um *score* final para cada solução (DEMŠAR, 2006).

Tabela 5.4: Teste de Friedman: ranqueamento dos métodos.

Ranqueamento	Método	Friedman's Score
1	HF+DL + (ReliefF+bGWO)	2,40
1	HF+DL + (ReliefF+GA)	2,40
2	HF + (ReliefF+PSO)	3,60
2	HF + (ReliefF+bGWO)	3,60
3	DL + (ReliefF+bGWO)	4,00
4	ResNet-50	6,20
5	EfficientNet-B2	6,80
6	DenseNet-121	7,20
7	VGG-19	8,60
8	Inception v3	9,40

Por fim, o *ensemble* heterogêneo de classificadores foi outro fator relevante para conquistar os resultados listados previamente. Conseguimos definir uma combinação de algoritmos que favoreceu o processo de reconhecimento de padrões de diferentes tipos de imagens H&E, com um sistema mais robusto e descentralizado, capaz de cobrir os pontos fracos que podem existir em um único classificador. Além disso, o *bias* e a variância foram reduzidos, minimizando a presença de *overfitting*.

5.2.1 Observações em relação aos trabalhos correlatos

Considerando os melhores resultados (Tabela 5.2 – *top 1*) da nossa proposta, uma visão geral é apresentada em relação à literatura, envolvendo cada conjunto H&E. É importante lembrar que, apesar de o tipo de imagem ser o mesmo, os *datasets* não são similares, seja no número de exemplos, classes, métricas utilizadas, validações ou amostras distintas. Logo, o intuito dessa verificação não é uma comparação direta entre as taxas de desempenhos, situação que exige ter condições de igualdade entre os modelos, por exemplo, reproduzindo cada estratégia nos conjuntos H&E aqui explorados. Esta tarefa é difícil ou mesmo inviável. Portanto, espera-se observar se as soluções obtidas fornecem resultados dentre os disponíveis na literatura especializada, conforme Tabelas 5.5 a 5.8 para os conjuntos CR, OED, LA e LG, e NHL.

A partir dos valores coletados para esse tipo de observação, nota-se que as soluções aqui obtidas conseguiram fornecer acurácias extremamente competitivas, especialmente em três (CR, OED e LG) das cinco bases H&E. No conjunto LA (Tabela 5.7), nosso método forneceu uma acurácia muito próxima das melhores soluções disponíveis na literatura: 99,24% (nossa solução) contra os desempenhos de 99,62% (ROBERTO et al., 2021) e 100% (NANNI et al., 2019). Em NHL, no entanto, mesmo que a solução tenha fornecido uma acurácia dentre as conquistadas nas propostas de (ANDREARCZYK; WHELAN, 2017; CANDELERO et al., 2020; ROBERTO et al., 2021), observa-se uma diferença mais bem definida em relação aos resultados de (NANNI et al., 2020; NANNI; GHIDONI; BRAHNAM, 2018; NANNI et al., 2019), mostrando que as estratégias aqui exploradas merecem mais atenção para este tipo de imagem. Apesar disso, essa visão geral permitiu ilustrar a capacidade discriminativa das soluções obtidas para diferentes tipos de imagens histológicas, considerando estratégias ainda não exploradas na literatura. Mais ainda, a metodologia aqui desenvolvida conseguiu fornecer soluções relevantes, com informações importantes sobre as melhores combinações de descritores e métodos de seleção, com a vantagem de utilizar um número reduzido de atributos e revelar suas principais ocorrências para reconhecer possíveis padrões em diferentes tipos de amostras histológicas.

Tabela 5.5: Visão geral dos valores de acurácias obtidos por diferentes abordagens para a classificação de imagens de câncer colorretal (CR).

Autor	Método	Acurácia
Método Proposto	DL + (ReliefF + bGWO)	100%
Roberto et al. (2021)	ResNet50 com <i>fine-tuning</i> , atributos fractais multidimensionais e multiescala	99,39%
Nanni, Ghidoni e Brahnham (2018)	8 modelos de CNN, descritores <i>handcrafted</i>	97,60%
Nanni et al. (2020)	9 modelos de CNN, descritores <i>handcrafted</i>	97,50%
Nanni et al. (2019)	6 modelos de CNN e descritores <i>handcrafted</i>	97,00%
Candelerio et al. (2020)	Le-Net, atributos fractais multidimensionais e multiescala, Haralick, LBP	91,06%

Tabela 5.6: Visão geral dos valores de acurácias obtidos por diferentes abordagens para a classificação de imagens de displasia oral (OED).

Autor	Método	Acurácia
Método Proposto	HF + (ReliefF + PSO)	100%
Azarmehr et al. (2022)	<i>Neural architecture search</i> e descritores <i>handcrafted</i> (morfológicos e não morfológicos)	95,20%
Adel et al. (2019)	Descritores <i>handcrafted</i> (SIFT, SURF, ORB)	92,80%
Silva et al. (2022)	Descritores <i>handcrafted</i> (morfológicos e não morfológicos)	92,40%

Tabela 5.7: Visão geral dos valores de acurácias obtidos por diferentes abordagens para a classificação de imagens de fígado (LA e LG).

Autor	Método	Acurácia (LA)	Acurácia (LG)
Nanni et al. (2019)	6 modelos de CNN e descritores <i>handcrafted</i>	100%	100%
Método Proposto	LA: HF + (ReliefF + bGWO) LG: HF + DL + (ReliefF + bGWO)	99,24%	100%
Roberto et al. (2021)	ResNet50 com <i>fine-tuning</i> , atributos fractais multidimensionais e multiescala	99,62%	99,62%
Andrearczyk e Whelan (2017)	CNN para textura	99,10%	98,20%
Watanabe, Kobayashi e Wada (2016)	Descritor <i>handcrafted</i> GIST	88,40%	93,70%

Tabela 5.8: Visão geral dos valores de acurácias obtidos por diferentes abordagens para a classificação de imagens de linfomas (NHL).

Autor	Método	Acurácia
Nanni et al. (2020)	9 modelos de CNN, descritores <i>handcrafted</i>	97,33%
Nanni, Ghidoni e Brahnam (2018)	8 modelos de CNN, descritores <i>handcrafted</i>	97,33%
Nanni et al. (2019)	6 modelos de CNN e descritores <i>handcrafted</i>	96,87%
Roberto et al. (2021)	ResNet50 com <i>fine-tuning</i> , atributos fractais multidimensionais e multiescala	95,55%
Método Proposto	HF + DL + (ReliefF + GA)	92,25%
Candelerio et al. (2020)	Le-Net, atributos fractais multidimensionais e multiescala, Haralick, LBP	82,01%
Andrearczyk e Whelan (2017)	CNN para textura	65,10%

Capítulo 6

Conclusão

Neste trabalho, foi proposto o *ensemble* de descritores *handcrafted* e *deep learned*, analisados mediante duplo estágio de seleção de atributos e classificados via *ensemble* de classificadores. O método foi aplicado para a tarefa de classificação de imagens histológicas H&E representativas de quatro categorias de imagens: câncer colorretal, tecidos hepáticos, tecidos orais e linfomas não-Hodgkin. Os descritores *handcrafted* foram obtidos via técnicas fractais multidimensionais e multiescala e atributos de textura LBP e Haralick. Quanto aos descritores *deep learned*, cinco arquiteturas CNN pré-treinadas no banco de imagens Imagenet, sem nenhum tipo de *fine-tuning*, foram utilizadas como extratores de atributos, sendo elas as arquiteturas ResNet-50, VGG-19, Inception v3, EfficientNet-B2 e DenseNet-121. O *ensemble* heterogêneo de classificadores foi composto pelos algoritmos SVM, KNN, *Random Forest*; Regressão Logística e *Naive Bayes*.

Os *ensembles* de descritores se mostraram superiores em relação às classificações somente por CNN, inclusive, em alguns casos, mais relevantes do que os conquistados com *ensembles* considerando somente *deep features*. A estratégia desenvolvida neste trabalho conseguiu fornecer valores de acurácias médias de 90,72% a 100%. Outra importante contribuição desse trabalho está relacionada com as principais soluções explorando um número reduzido de atributos. Considerando que o espaço inicial de

características é 11086 valores (462 descritores *handcrafted* com 10624 *deep learned*), a metodologia proposta conseguiu fornecer taxas de acurácias relevantes usando no máximo 53 características. Os destaques foram: CR com apenas 22 descritores via bGWO; OED com apenas 11 descritores via PSO; LA com 39 descritores via bGWO; LG com somente 29 atributos, bGWO; NHL com 53 descritores, GA.

Além disso, a metodologia proposta atingiu resultados competitivos em relação aos trabalhos que abordaram os mesmos contextos de imagens, fornecendo as principais combinações de técnicas para apoiar o desenvolvimento de sistemas CAD focados em analisar imagens histológicas H&E, especialmente representativas de câncer colorretal, tecido hepático, linfomas não-Hodgkin e displasia oral.

Em relação às limitações ou restrições enfrentadas pela aplicação do método, podemos elencar principalmente a questão dos algoritmos que utilizam algum tipo de hiperparâmetros. Uma vez que foram utilizados valores fixos de hiperparâmetros, pode haver a possibilidade deles não representarem os valores ideais na solução do contexto proposto, já que não foi realizada qualquer otimização neste quesito. Tais hiperparâmetros abrangem os métodos de extração de atributos *handcrafted* e os métodos de seleção de atributos: ponto de corte do algoritmo ReliefF e valores intrínsecos para a execução dos seletores *wrapper*.

Por fim, nosso trabalho proveu as seguintes contribuições: a construção e análise comparativa de *ensembles* baseados em múltiplas categorias de descritores *handcrafted* e *deep learned*; estruturação e aplicação do método de *ensemble learning* ainda não explorado no contexto H&E, composto por um duplo estágio de seleção de atributos e um *ensemble* heterogêneo de classificadores; comparação do método proposto frente às arquiteturas CNN treinadas para a tarefa de classificação das imagens em análise; conhecer alguns dos padrões de técnicas e atributos presentes nas melhores soluções para a classificação de imagens histológicas representativas do câncer colorretal, displasia oral, linfomas não-Hodgkin e tecidos hepáticos; obtenção de acurácias

competitivas frente aos trabalhos correlatos, com a vantagem de explorar um número reduzido de atributos.

6.1 Trabalhos Futuros

Visando possíveis aprimoramentos do trabalho apresentado, outros aspectos podem ser explorados. Um deles é a otimização dos hiperparâmetros utilizados pelos seletores *wrapper*, tais como número da população, número de iterações, probabilidade de mutação, entre outros. Além disso, dos hiperparâmetros utilizados na extração dos atributos *handcrafted*: tamanho L do hipercubo para os métodos multidimensionais e multiescala; e do comprimento do vetor deslizante d no método de Higuchi. É válido ampliar o escopo das arquiteturas de redes neurais ou explorar outras arquiteturas que não as CNN, como, por exemplo, a pertinência do uso dos *Vision Transformers* (DOSOVITSKIY et al., 2020). A exploração de diferentes formas de concatenar os atributos extraídos e de construir o *ensemble* de classificadores também é válido.

6.2 Contribuições

O desenvolvimento deste trabalho permitiu publicações relacionadas direta e indiretamente com o tema, a saber:

- Artigo: *Multidimensional and multiscale higuchi dimension for the analysis of colorectal histological images*. Em *2020 IEEE International Conference on Bioinformatics and Biomedicine* (BIBM). IEEE, p. 2833–2839, 2020 (TENGUAM et al., 2020).
- Artigo: *Classification of H&E images exploring ensemble learning with two-stage feature selection*. Em *29th International Conference on Systems, Signals and Image Processing* (IWSSIP). IEEE, p.1-4, 2022 (TENGUAM et al., 2022).

Referências

ADEL, D.; MOUNIR, J.; EL-SHAFFEY, M.; ELDIN, Y. A.; MASRY, N. E.; ABDEL-RAOUF, A.; ELHAMID, I. S. A. Oral epithelial dysplasia computer aided diagnostic approach. In: . [S.l.]: Institute of Electrical and Electronics Engineers Inc., 2019. p. 313–318.

AGEMAP, N. I. o. A. The atlas of gene expression in mouse aging project (agemap). 2020. <https://ome.grc.nia.nih.gov/iicbu2008/agemap/index.html>. Online; acesso: 19-03-2020.

AGGARWAL, C. C. Data Classification: Algorithms and Applications. [S.l.]: Chapman & Hall/CRC, 2015. v. 1.

AHA, D. W. Lazy Learning. [S.l.]: Springer, 1997. v. 11.

AHMED, T.; SABAB, N. H. N. Classification and understanding of cloud structures via satellite images with efficientunet. 9 2020. Disponível em: <http://arxiv.org/abs/2009.12931>.

AHONEN, T.; HADID, A.; PIETIKÄINEN, M. Face recognition with local binary patterns. ECCV 2004: Computer Vision - ECCV 2004, Springer, v. 3021, p. 469–481, 2004.

AKBAR, B.; GOPI, V. P.; BABU, V. S. Colon cancer detection based on structural and statistical pattern recognition. 2015 2nd International Conference on Electronics and Communication Systems (ICECS), IEEE, p. 1735—1739, 2015.

AL-MASNI, M. A.; AL-ANTARI, M. A.; PARK, J. M.; GI, G.; KIM, T. Y.; RIVERA, P.; VALAREZO, E.; HAN, S. M.; KIM, T. S. Detection and classification of the breast abnormalities in digital mammograms via regional convolutional neural network. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, p. 1230–1233, 2017.

ALMARAZ-DAMIAN, J.-A.; PONOMARYOV, V.; SADOVNYCHYI, S.; CASTILLEJOS-FERNANDEZ, H. Melanoma and nevus skin lesion classification using handcraft and deep learning feature fusion via mutual information measures. Entropy, v. 22, n. 4, p. 484, 2020.

ALTUNBAY, D.; CIGIR, C.; SOKMENSUER, C.; GUNDUZ-DEMIR, C. Color graphs for automated cancer diagnosis and grading. IEEE Transactions on Biomedical Engineering, v. 57, n. 3, p. 665–674, 2010.

ANDREARCZYK, V.; WHELAN, P. F. Deep learning for biomedical texture image analysis. Proceedings of the Irish Machine Vision & Image Processing Conference. Irish Pattern Recognition & Classification Society (IPRCS), 2017.

ARAÚJO, T.; ARESTA, G.; CASTRO, E.; ROUCO, J.; AGUIAR, P.; ELOY, C.; POLÓNIA, A.; CAMPILHO, A. Classification of breast cancer histology images using convolutional neural networks. PLoS One, v. 12, n. 6, p. e0177544, 2017.

AZARMEHR, N.; SHEPHARD, A.; MAHMOOD, H.; RAJPOOT, N.; KHURRAM, S. A. Automated oral epithelial dysplasia grading using neural networks and feature analysis. In: Medical Imaging with Deep Learning. [s.n.], 2022. Disponível em: <https://openreview.net/forum?id=ABl-dIO4g74>.

BAISH, J. W.; JAIN, R. K. Fractals and cancer. Cancer research, v. 60, n. 14, p. 3683–3688, 2000.

BEN HAMIDA, A.; DEVANNE, M.; WEBER, J.; TRUNTZER, C.; DERANGÈRE, V.; GHIRINGHELLI, F.; FORESTIER, G.; WEMMERT, C. Deep learning for colon cancer histopathological images analysis. Computers in Biology and Medicine, v. 136, p. 104730, 2021. ISSN 0010-4825. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0010482521005242>.

BHOWAL, P.; SEN, S.; SARKAR, R. A two-tier feature selection method using coalition game and nystrom sampling for screening covid-19 from chest x-ray images. Journal of Ambient Intelligence and Humanized Computing, 2021. Disponível em: <https://doi.org/10.1007/s12652-021-03491-4>.

BOUSSAÏD, I.; LEPAGNOT, J.; SIARRY, P. A survey on optimization metaheuristics. Information Sciences, Elsevier, v. 237, p. 82–117, 2013.

BRADLEY, A. P. The use of the area under the roc curve in the evaluation of machine learning algorithms. Pattern Recognition, v. 30, n. 7, p. 1145–1159, 1997.

BREIMAN, L. Random forests. Machine Learning, v. 45, n. 1, p. 5–32, 2001.

BROADBENT, S. R.; HAMMERSLEY, J. M. Percolation processes: I. crystals and mazes. Mathematical Proceedings of the Cambridge Philosophical Society, Cambridge University Press, v. 53, n. 3, p. 629–641, 1957.

BUDNIK, M.; GUTIERREZ-GOMEZ, E.; SAFADI, B.; PELLERIN, D.; QUÉNOT, G. Learned features versus engineered features for multimedia indexing. Multimedia Tools and Applications, Springer, v. 76, p. 11941–11958, 2017.

BUNDE, A.; HAVLIN, S. A brief introduction to fractal geometry. Fractals in Science, Springer Berlin Heidelberg, 1994.

CANDELERO, D.; ROBERTO, G. F.; NASCIMENTO, M. Z. do; ROZENDO, G. B.; NEVES, L. A. Selection of cnn, haralick and fractal features based on evolutionary algorithms for classification of histological images. In: 2020 IEEE International

Conference on Bioinformatics and Biomedicine (BIBM). [S.l.]: IEEE, 2020. p. 2709–2716.

CARUSO, D.; ZERUNIAN, M.; CIOLINA, M.; DE SANTIS, D.; RENGO, M.; SOMMO, M. H.; GIUNTA, G.; CONFORTO, S.; SCHMID, M.; NERI, E.; LAGHI, A. Haralick's texture features for the prediction of response to therapy in colorectal cancer: a preliminary study. La radiologia medica, Springer, v. 123, p. 161–167, 2018.

CHAN, H. P.; HADJIISKI, L. M.; SAMALA, R. K. Computer-aided diagnosis in the era of deep learning. Med. Phys., v. 47, n. 5, p. e218–e227, 2020.

CHANDRASHEKAR, G.; SAHIN, F. A survey on feature selection methods. Computers and Electrical Engineering, v. 40, n. 1, p. 16–28, 2014.

CONNERS, R. W.; TRIVEDI, M. M.; HARLOW, C. A. Segmentation of a high-resolution urban scene using texture operators. Comp Vision Graphics Image Process, Elsevier, v. 25, n. 3, p. 273–310, 1984.

CUEVAS, E.; GONZALEZ, M.; ZALDIVAR, D.; PEREZ-CISNEROS, M.; GARCÍA, G. An algorithm for global optimization inspired by collective animal behavior. Discrete Dynamics in Nature and Society, v. 2012, p. 24, 2012.

CĂLIMAN A. ; IVANOVICI, M. Psoriasis image analysis using color lacunarity. 2012 13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), IEEE, p. 1401–1406, 2012.

DEMŠAR, J. Statistical comparisons of classifiers over multiple data sets. Journal of Machine Learning Research, v. 7, p. 1–30, 2006. ISSN 15337928.

DEVOS, A. K.; van Huffel, S.; SIMONETTI, A. W.; GRAAF, M. v.; HEERSCHAP, A.; BUYDENS, L. M. C. Chapter 11 - classification of brain tumours by pattern recognition of magnetic resonance imaging and spectroscopic data. In: TAKTAK, A. F. G.; FISHER, A. C. (Ed.). Outcome Prediction in Cancer. [S.l.]: Elsevier, 2007. p. 285 – 318.

DORNAIKA, F.; BOSAGHZADEH, A.; SALMANE, H.; RUICHEK, Y. Chapter 9 - object categorization using adaptive graph-based semi-supervised learning. In: SAMUI, P.; SEKHAR, S.; BALAS, V. E. (Ed.). Handbook of Neural Computation. [S.l.]: Academic Press, 2017. p. 167–179. ISBN 978-0-12-811318-9.

DOSOVITSKIY, A.; BEYER, L.; KOLESNIKOV, A.; WEISSENBORN, D.; ZHAI, X.; UNTERTHINER, T.; DEHGhani, M.; MINDERER, M.; HEIGOLD, G.; GELLY, S.; USZKOREIT, J.; HOULSBY, N. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. 2020. Disponível em: <http://arxiv.org/abs/2010.11929>.

EGER, S.; YOUSSEF, P.; GUREVYCH, I. Is it time to swish? comparing deep learning activation functions across nlp tasks. In: . [S.l.: s.n.], 2018. p. 4415–4424.

ELGENDY, M. Advanced CNN Architectures. 2020. <https://livebook.manning.com/book/grokking-deep-learning-for-computer-vision/chapter-5/v-3/34>. Online; acesso: 30-10-2020.

ELMANNAI, H.; HAMDI, M.; ALGARNI, A. Deep learning models combining for breast cancer histopathology image classification. International Journal of Computational Intelligence Systems, v. 14, n. 1, p. 1003–1013, 2021. ISSN 18756883.

EMARYA, E.; ZAWBAAB, H. M.; HASSANIENA, A. E. Binary grey wolf optimization approaches for feature selection. Neurocomputing, Elsevier, p. 371–381, 2016.

FAWCETT, F. An introduction to roc analysis. Pattern Recognition Letters, v. 27, n. 8, p. 861–874, 2006.

FONDÓN, I.; SARMIENTO, A.; GARCÍA, A. I.; SILVESTRE, M.; ELOY, C.; POLÓNIA, A.; AGUIAR, P. Automatic classification of tissue malignancy for breast carcinoma diagnosis. Computers in Biology and Medicine, Elsevier, v. 96, n. 1, p. 41–51, 2018.

FONTES, L. R. G. Notas em percolação. Instituto de Matemática Pura e Aplicada, 1996.

FREUND, Y.; SCHAPIRE, R. E. Experiments with a new boosting algorithm. Proceedings of the 13th International Conference on Machine Learning, p. 148–156, 1996.

GARTNER, L. P. Textbook of Histology. 5. ed. [S.l.]: Elsevier, 2020.

GHOSH, S.; BANDYOPADHYAY, A.; SAHAY, S.; GHOSH, R.; KUNDU, I.; SANTOSH, K. C. Colorectal histology tumor detection using ensemble deep neural network. Engineering Applications of Artificial Intelligence, v. 100, p. 104202, 2021.

GOMES, R. L.; VANDERLEI, L. C. M.; GARNER, D. M.; VANDERLEI, F. M.; VALENTI, V. E. Higuchi fractal analysis of heart rate variability is sensitive during recovery from exercise in physically active men. MedicalExpress, v. 4, n. 3, 2017.

GOMOLKA, R. S.; KAMPUSCH, S.; KANIUSAS, E.; THÜRK F. SZÉLES, J. C.; KLONOWSKI, W. Higuchi fractal dimension of heart rate variability during percutaneous auricular vagus nerve stimulation in healthy and diabetic subjects. Front. Physiol., v. 9, n. 1162, 2018.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep learning. MIT Press, 2016.

GOOGLE. Colaboratory. 2021. Online; acesso: 20-07-2021. Disponível em: <https://research.google.com/colaboratory/faq.html>.

GRANIK, M.; MESYURA, V. Fake news detection using naive bayes classifier. In: 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON). [S.l.: s.n.], 2017. p. 900–903.

GREENSPAN, H.; GINNEKEN, B.; SUMMERS, R. M. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. IEEE Transactions on Medical Imaging, IEEE, v. 35, n. 5, p. 1153 – 1159, 2016.

HAGERTY, J. R.; STANLEY, R. J.; ALMUBARAK, H. A.; LAMA, N.; KASMI, R.; GUO, P.; DRUGGE, R. J.; RABINOVITZ, H. S.; OLIVIERO, M.; STOECKER, W. V. Deep learning and handcrafted method fusion: Higher diagnostic accuracy for melanoma dermoscopy images. IEEE Journal of Biomedical and Health Informatics, IEEE, v. 23, n. 4, p. 1385–1391, 2019.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The weka data mining software: an update. ACM SIGKDD explorations newsletter, ACM, v. 11, n. 1, p. 10–18, 2009.

HAQUE, I. R. I.; NEUBERT, J. Deep learning approaches to biomedical image segmentation. Informatics in Medicine Unlocked, Elsevier, v. 18, p. 100297, 2020.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics, IEEE, SMC-3, n. 6, p. 610 – 621, 1973.

HARRIS, J. W.; STOCKER, H. Handbook of Mathematics and Computational Science. [S.l.]: Springer, 1998.

HASSAN, A. H.; WAHED, M. E.; ATIEA, M. A.; METWALLY, M. S. A hybrid approach for classification breast cancer histopathology images. Frontiers in Scientific Research and Technology, Suez University, Faculty of Science, v. 3, n. 1, p. 1–10, 2022. ISSN 2682-2962.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Proceedings of the IEEE International Conference on Computer Vision (ICCV), IEEE, p. 1026–1034, 2015.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, p. 770–778, 2016.

HIGUCHI, T. Approach to an irregular time series on the basis of the fractal theory. Physica D: Nonlinear Phenomena, Elsevier, v. 31, n. 2, p. 227–283, 1988.

HSU, H.; HSIEH, C. W.; LU, M. Hybrid feature selection by combining filters and wrappers. Expert Systems with Applications, v. 38, n. 7, p. 8144–8150, 2011.

HUANG, G.; LIU, Z.; MAATEN, L.; WEINBERGER, K. Q. Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 2261–2269, 2017.

IDE, H.; KURITA, T. Improvement of learning for cnn with relu activation by sparse regularization. 2017 International Joint Conference on Neural Networks (IJCNN), IEEE, p. 2684–2691, 2017.

IFTIKHAR, A. Feature selection using particle swarm optimization in intrusion detection. International Journal of Distributed Sensor Networks, v. 2015, p. 1–8, 2015.

IFTIKHAR, M. A.; HASSAN, M.; ALQUHAYZ, H. A colon cancer grade prediction model using texture and statistical features, smote and mrmr. 2016 19th International Multi-Topic Conference (INMIC), IEEE, p. 1–7, 2016.

IVANOVICI, M.; RICHARD, N. The lacunarity of colour fractal images. In: Proceedings of the 16th IEEE international conference on image processing (ICIP), IEEE, p. 453–456, 2009.

IVANOVICI, M.; RICHARD, N. Fractal dimension of color fractal images, image processing. IEEE Transactions on Image Processing, IEEE, v. 20, n. 1, p. 227–235, 2011.

IVANOVICI, M.; RICHARD, N.; DECEAN, H. Fractal dimension and lacunarity of psoriatic lesions—a colour approach—. Proceedings of the 2nd WSEAS International Conference on Biomedical Electronics and Biomedical Informatics, BEBI '09, v. 6, 2009.

JASTI, V. D. P.; ZAMANI, A. S.; ARUMUGAM, K.; NAVED, M.; PALLATHADKA, H.; SAMMY, F.; RAGHUVANSHI, A.; KALIYAPERUMAL, K. Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis. Security and Communication Networks, 2022. Disponível em: <https://doi.org/10.1155/2022/1918379>.

JIANG, L.; CAI, Z.; WANG, D.; JIANG, S. Survey of improving k-nearest-neighbor for classification. In: Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007). [S.l.: s.n.], 2007. v. 1, p. 679–683.

JUNQUEIRA, L. C.; CARNEIRO, J. Histologia Básica: Texto & Atlas. 12. ed. [S.l.]: Guanabara Koogan, 2013.

JØRGENSEN, A.; RASMUSSEN, A. M.; ANDERSEN, N. K. M.; ANDERSEN, S. K.; EMBORG, J.; RØGE, R.; ØSTERGAARD, L. Using cell nuclei features to detect colon cancer tissue in hematoxylin and eosin stained slides. Cytometry Part A, Wiley Online Library, 2017.

KALRA, S.; TIZHOOSH, H. R.; CHOI, C.; SHAH, S.; DIAMANDIS, P.; CAMPBELL, C. J. V.; PANTANOWITZ, L. Yottixel – an image search engine for large archives of histopathology whole slide images. Medical Image Analysis, v. 65, p. 101757, 2020. ISSN 1361-8415.

KANTARCIOGLU, M.; VAIDYA, J. Privacy-preserving naive bayes classifier for horizontally partitioned data. IEEE Workshop on Privacy-Preserving Data Mining, IEEE, 2003.

KARSMAKERS, P.; PELCKMANS, K.; SUYKENS, J. A. K. Multi-class kernel logistic regression: a fixed-size implementation. In: 2007 International Joint Conference on Neural Networks. [S.l.: s.n.], 2007. p. 1756–1761.

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. CASCON '19: Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering, p. 92–99, 2019.

KATHER, J. N.; WEIS, C.; BIANCONI, F.; MELCHERS, S. M.; SCHAD, L. R.; GAISER, T.; MARX, A.; ZOELLNER, F. G. Multi-class texture analysis in colorectal cancer histology. Scientific Reports, v. 6, 2016.

KENNEDY, J.; EBERHART, R. Particle swarm optimization. Proceedings of ICNN'95 - International Conference on Neural Networks, IEEE, v. 4, p. 1942–1948, 1995.

KŁONOWSKI, W.; PIERZCHALSKI, M.; STEPIEN, P.; STEPIEN, R.; AHAMMER, H. Application of higuchi's fractal dimension in analysis of images of anal intraepithelial neoplasia. Chaos, Solitons & Fractals, Elsevier, v. 48, p. 54–60, 2013.

KŁONOWSKI, W.; STEPIEN, P.; STEPIEN, R.; SEDIVY, R.; AHAMMER, H.; SPASIC, S. Analysis of anal intraepithelial neoplasia images using 1d and 2d higuchi's fractal dimension methods. Fractals, World Scientific Publishing Company, v. 26, n. 3, 2018.

KORKMAZ, S. A.; BINOL, H. Classification of molecular structure images by using ann, rf, lbp, hog, and size reduction methods for early stomach cancer detection. Journal of Molecular Structure, v. 1156, p. 255–263, 2018.

KOTSIANTIS, S. B. Feature selection for machine learning classification problems: a recent overview. Artificial Intelligence Review, Springer, v. 41, 2011.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F.; BURGESS, C. J. C.; BOTTOU, L.; WEINBERGER, K. Q. (Ed.). Advances in Neural Information Processing Systems 25. [S.l.]: Curran Associates, Inc., 2012. p. 1097–1105.

KULKARNI, U.; MEENA, S. M.; GURLAHOSUR, S. V.; MUDENGUDI, U. Classification of cultural heritage sites using transfer learning. 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), IEEE, p. 391–397, 2019.

KUMAR, A.; BI, L.; KIM, J.; FENG, D. D. Chapter five - machine learning in medical imaging. In: FENG, D. D. (Ed.). Biomedical Information Technology. 2. ed. [S.l.]: Elsevier, 2020. p. 167–196.

KUMAR, N.; SHARMA, M.; SINGH, V. P.; MADAN, C.; MEHANDIA, S. An empirical study of handcrafted and dense feature extraction techniques for lung and colon cancer classification from histopathological images. Biomedical Signal Processing and Control, Elsevier Ltd, v. 75, n. 75, p. 103596, 2022. ISSN 17468108. Disponível em: <https://doi.org/10.1016/j.bspc.2022.103596>.

KUNCHEVA, L. I. Combining classifiers: soft computing solutions. In: Pattern Recognition. [S.l.: s.n.], 2001. p. 427–451.

KUNCHEVA, L. I.; BEZDEK, J. C.; DUIN, R. P. W. Decision templates for multiple classifier fusion: an experimental comparison. Pattern Recognition, v. 34, n. 2, p. 299 – 314, 2001.

KURT, I.; TURE, M.; KURUM, A. T. Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. Expert Systems with Applications, v. 34, n. 1, p. 366 – 374, 2008.

KWASIGROCH, A.; MIKOŁAJCZYK, A.; GROCHOWSKI, M. Deep neural networks approach to skin lesions classification — a comparative analysis. In: 2017 22nd International Conference on Methods and Models in Automation and Robotics (MMAR). [S.l.]: IEEE, 2017. p. 1069–1074.

LAHDENOJA, O.; POIKONEN J. ANTD LAIHO, M. Towards understanding the formation of uniform local binary patterns. ISRN Machine Vision, v. 2013, 07 2013.

LAWRENCE, S.; GILES, C. L.; TSOI, A. C.; BACK, A. D. Face recognition: A convolutional neural-network approach. IEEE Transactions on Neural Networks, IEEE, v. 8, n. 1, p. 98–113, 1997.

LE, T. T.; URBANOWICZ, R. J.; MOORE, J. H.; MCKINNEY, B. A. Statistical inference relief (stir) feature selection. Bioinformatics, v. 35, n. 8, p. 1358–1365, 2019.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. Nature, Macmillan Publishers Limited, v. 521, n. 7553, p. 436–444, 2015.

LI, L.; CHANG, L.; KE, S.; HUANG, D. Multifractal analysis and lacunarity analysis: A promising method for the automated assessment of muskmelon (cucumis melol.) epidermis netting. Computers and Electronics in Agriculture, Elsevier, v. 88, p. 72–84, 2012.

LI, X.; SHEN, X.; ZHOU, Y.; WANG, X.; LI, T. Q. Classification of breast cancer histopathological images using interleaved densenet with senet (idsnet). PLoS ONE, Public Library of Science, v. 15, 5 2020.

LI, X.; WU, J.; JIANG, H.; CHEN, E. Z.; DONG, X.; RONG, R. Skin lesion classification via combining deep learning features and clinical criteria representations. bioRxiv, Cold Spring Harbor Laboratory, 2018.

LI, Y.; SHEN, Y.; FAN, X.; HUANG, X.; YU, H.; ZHAO, G.; MA, W. A novel EEG-based major depressive disorder detection framework with two-stage feature selection. BMC Medical Informatics and Decision Making, v. 22, n. 209, 2022. Disponível em: <https://doi.org/10.1186/s12911-022-01956-w>.

LIEPINS, G. E.; VOSE, M. D. Representational issues in genetic optimization. Journal of Experimental & Theoretical Artificial Intelligence, Taylor & Francis, v. 2, n. 2, p. 101–115, 1990.

LITJENS, G.; KOOL, T.; BEJNORDI, B. E.; SETIO, A. A. A.; CIOMPI, F.; GHAFORIAN, M.; LAAK, J. A. W. M. van der; GINNEKEN B. SÁNCHEZ, C. I. A survey on deep learning in medical image analysis. Medical Image Analysis, Elsevier, v. 42, p. 60–88, 2017.

LOPES, R.; BETROUNI, N. Fractal and multifractal analysis: a review. Medical image analysis, Elsevier, v. 13, n. 4, p. 634–649, 2009.

MAHMOOD, A.; GIRALDO, A.; BENNAMOUN, M.; AN, S.; SOHEL, F.; BOUSSAID, F.; HOVEY, R.; FISHER, R.; KENDRICK, G. Automatic hierarchical classification of kelps using deep residual features. Sensors, v. 20, p. 447, 01 2020.

MANDELBROT, B. The Fractal Geometry of Nature. [S.l.]: Freeman, 1983.

MANDELBROT, B. B. Fractals: Form, chance and dimension. M. San Francisco: Freeman, 1977.

MATHWORKS. R2019a at a Glance. 2021. Online; acesso: 20-07-2021. Disponível em: <https://ch.mathworks.com/solutions/deep-learning/models.html>.

MATHWORKS. Extract local binary pattern (LBP) features. c2022. <https://www.mathworks.com/help/vision/ref/extractlbpfeatures.html#buumhti-1-CellSize>. Online; acesso: 18-05-2022.

MAZO, C.; ALEGRE, E.; TRUJILLO, M. Classification of cardiovascular tissues using lbp based descriptors and a cascade svm. Computer Methods and Programs in Biomedicine, Elsevier, v. 147, p. 1–10, 2017.

MENGDI, L.; LIANCHENG, X.; JING, Y.; JIE, H. A feature gene selection method based on relieff and pso. In: 2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA). [S.l.]: IEEE, 2018. p. 298–301.

MIYAMOTO, Y.; MERRYMAN, T. E. Fast calculation of haralick texture features. Technical Report, 2005.

MUNIEN, C.; VIRIRI, S. Classification of hematoxylin and eosin-stained breast cancer histology microscopy images using transfer learning with efficientnets. Computational Intelligence and Neuroscience, v. 2021, p. 1–17, 04 2021.

NANNI, L.; BRAHNAM, S.; GHIDONI, S.; MAGUOLO, G. General purpose (genp) bioimage ensemble of handcrafted and learned features with data augmentation. CoRR, abs/1904.08084, 2019.

NANNI, L.; GHIDONI, S.; BRAHNAM, S. Handcrafted vs. non-handcrafted features for computer vision classification. Pattern Recognition, Elsevier, v. 71, p. 158–172, 2017.

NANNI, L.; GHIDONI, S.; BRAHNAM, S. Ensemble of convolutional neural networks for bioimage classification. Applied Computing and Informatics, Elsevier, 2018.

NANNI, L.; GHIDONI, S.; BRAHNAM, S.; LIU, S.; ZHANG, L. Ensemble of hand-crafted and deep learned features for cervical cell classification. In: NANNI, L.; BRAHNAM, S.; BRATTIN, R.; GHIDONI, S.; JAIN, L. (Ed.). Deep Learners and Deep Learner Descriptors for Medical Applications. Intelligent Systems Reference Library. [S.l.]: Springer, 2020. v. 186, p. 117–135.

NAYAK, S. R.; MISHRA, J. An improved method to estimate the fractal dimension of colour images. Perspectives in Science, Elsevier, v. 8, p. 412–416, 2016.

NEVES, L. A.; NASCIMENTO, M. do; OLIVEIRA, D. L.; MARTINS, A. S.; GODOY, M. F.; ARRUDA, P. F.; NETO, D. d. S.; MACHADO, J. M. 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.

NGUYEN, H.-T.; VU, N. S.; CAPLIER, A. How far we can improve micro features based face recognition systems? In: . [S.l.: s.n.], 2012. p. 350–353.

NIKOLAIDIS, N. S.; NIKOLAIDIS, I. N.; TSOUROS, C. C. A variation of the box-counting algorithm applied to colour images. arXiv preprint arXiv:1107.2336, 2011.

NKETIAH, G.; ELSCHOT, M.; KIM, E.; TERUEL, J. R.; SCHEENEN, T. W.; BATHEN, T. F.; SELNÆS, K. M. T2-weighted mri-derived textural features reflect prostate cancer aggressiveness: preliminary results. Eur Radiol, Springer, v. 27, p. 3050–3059, 2017.

NWANKPA, C.; IJOMAH, W.; GACHAGAN, A.; MARSHALL, S. Activation functions: Comparison of trends in practice and research for deep learning. ArXiv, abs/1811.03378, 2018.

OHATA, E. F.; DAS CHAGAS, J. V. S.; BEZERRA, G. M.; HASSAN, M. M.; DE ALBUQUERQUE, V. H. C.; FILHO, P. P. R. A novel transfer learning approach for the classification of histological images of colorectal cancer. The Journal of Supercomputing, Springer US, v. 77, p. 9494–9519, 2021. ISSN 15730484.

OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. Pattern Recognition, v. 29, n. 1, p. 51–59, 1996.

PACHECO, M. A. C. Algoritmos Genéticos: Princípios e Aplicações. Apostila de Aula. Versão 1. Pontifícia Universidade Católica do Rio de Janeiro, ICA: Laboratório de Inteligência Computacional Aplicada. 1999. <http://www2.ica.ele.puc-rio.br/Downloads/38/CE-Apostila-Comp-Evol.pdf/>. Online; acesso: 2019-04-11.

PAN, S. J.; YANG, Q. A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering, v. 22, n. 10, p. 1345–1359, 2010.

PARMANTO, B. A.; MUNRO, P. W.; DOYLE, H. R. Improving committee diagnosis with resampling techniques. In: TOURETZKY, D. S.; MOZER, M. C.; HASSELMO, M. E. (Ed.). Advances in Neural Information Processing Systems. [S.l.]: MIT Press, 1996. p. 882–888.

PEDRINI, H.; SCHWARTZ, W. R. Análise de imagens digitais: Princípios, algoritmos e aplicações. 1. ed. [S.l.]: Cengage, 2008. v. 1.

PENTLAND, A. Fractal-based description of natural scenes. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6, n. 6, p. 661–674, 1984.

PLOTNICK, R. E.; GARDNER, R. H.; O'NEIL, R. V. Lacunarity indices as measures of landscape texture. Landscape Ecology, SPB Academic Publishing, v. 8, n. 3, p. 201–211, 1993.

PONTI JR., M. P. Combining classifiers: From the creation of ensembles to the decision fusion. In: 2011 24th SIBGRAPI Conference on Graphics, Patterns, and Images Tutorials. [S.l.: s.n.], 2011. p. 1–10.

PONTI, M. P.; PAPA, J. P. Improving accuracy and speed of optimum-path forest classifier using combination of disjoint training subsets. In: SANSONE, C.; KITTLER, J.; ROLI, F. (Ed.). Multiple Classifier Systems. [S.l.]: Springer Berlin Heidelberg, 2011. p. 237–248.

PYTORCH. Pytorch. 2021. Online; acesso: 20-07-2021. Disponível em: <https://pytorch.org/>.

PYTORCH. Models and Pre-trained Weights. 2022. Online; acesso: 25-09-2022. Disponível em: <https://pytorch.org/vision/stable/models.html>.

RADWAN, N. Leveraging sparse and dense features for reliable state estimation in urban environments. In: . [S.l.: s.n.], 2019.

RAJINIKANTH, V.; RAJ, A. J.; THANARAJ, K.; NAIK, G. A customized vgg19 network with concatenation of deep and handcrafted features for brain tumor detection. Appl. Sci., v. 10, n. 10, p. 3429, 2020.

RAKHLIN, A.; SHVETS, A.; IGLOVIKOV, V.; KALININ, A. A. Deep convolutional neural networks for breast cancer histology image analysis. Image Analysis and Recognition. ICIAR 2018, Springer, v. 10882, p. 737–744, 2018.

RAMANESWARAN, S.; SRINIVASAN, K.; VINCENT, P. M. D. R.; CHANG, C. Hybrid inception v3 xgboost model for acute lymphoblastic leukemia classification. Computational and Mathematical Methods in Medicine, v. 2021, p. 1–10, 07 2021.

REZVANTALAB, A.; SAFIGHOLI, H.; KARIMIJEHNI, S. Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms. arXiv:1810.10348, 2018.

RIBEIRO, M. G.; NEVES, L. A.; NASCIMENTO, M. Z. d.; ROBERTO, G. F.; MARTINS, A. M.; TOSTA, T. A. A. Classification of colorectal cancer based on the association of multidimensional and multiresolution features. Expert Systems With Applications, Elsevier, v. 120, p. 262—278, 2019.

ROBERTO, G. F.; LUMINI, A.; NEVES, L. A.; NASCIMENTO, M. Z. Fractal neural network: A new ensemble of fractal geometry and convolutional neural networks for the classification of histology images. Expert Systems With Applications, Elsevier, v. 166, p. 114103, 2021.

ROBERTO, G. F.; NASCIMENTO, M. Z.; MARTINS, A. S.; TOSTA, T. A. A.; FARIA, P. R.; NEVES, L. A. Classification of breast and colorectal tumors based on percolation of color normalized images. Computers & Graphics, Elsevier, v. 84, p. 134–143, 2019.

ROBERTO, G. F.; NEVES, L. A.; NASCIMENTO, M. Z.; TOSTA, T. A. A.; LONGO, L. C.; MARTINS, A. S.; FARIA, P. R. Features based on the percolation theory for quantification of non-hodgkin lymphomas. Computers in Biology and Medicine, Elsevier, v. 91, p. 135–147, 2017.

ROBNIK ŠIKONJA, M.; KONONENKO, I. Theoretical and empirical analysis of relief and rrelief. Machine Learning, Kluwer Academic Publishers, v. 53, p. 23–69, 2003.

RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA S. HUANG, Z. Imagenet large scale visual recognition challenge. International Journal of Computer Vision, Springer, v. 115, p. 211–252, 2015.

SANTOS, F. P. dos; PONTI, M. S. Alignment of local and global features from multiple layers of convolutional neural network for image classification. In: 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). [S.l.]: IEEE, 2019. p. 241–248.

SEN, S.; SAHA, S.; CHATTERJEE, S.; MIRJALILI, S.; SARKAR, R. A bi-stage feature selection approach for covid-19 prediction using chest ct images. Applied Intelligence, v. 51, p. 8985–9000, 2021. Disponível em: <https://doi.org/10.1007/s10489-021-02292-8>.

SETHY, P. K.; BEHERA, S. K. Automatic classification with concatenation of deep and handcrafted features of histological images for breast carcinoma diagnosis. Multimedia Tools and Applications, Springer US, v. 81, p. 9631–9643, 2022.

SHAMIR, L.; ORLOV, N.; ECKLEY, D. M.; MACURA, T. J.; GOLDBERG, I. G. IICBU 2008: a proposed benchmark suite for biological image analysis. Med Biol Eng Comput, Springer, v. 46, n. 9, p. 943–947, 2008.

SHARMA, S.; KUMAR, S. The Xception model: A potential feature extractor in breast cancer histology images classification. ICT Express, Elsevier B.V., v. 8, n. 1, p.

101–108, 2022. ISSN 24059595. Disponível em: <https://doi.org/10.1016/j.ict.2021.11.010>.

SHEN, D.; WU, G.; SUK, H. Deep learning in medical image analysis. Annual Review of Biomedical Engineering, v. 19, n. 1, p. 221–248, 2017.

SIEGEL, R. L.; MILLER, K. D.; FUCHS, H. E.; JEMAL, A. Cancer statistics, 2021. CA: a cancer journal for clinicians, v. 71, n. 1, p. 7–33, 2021.

SILVA, A. B.; MARTINS, A. S.; TOSTA, T. A. A.; NEVES, L. A.; SERVATO, J. P. S.; de Araújo, M. S.; DE FARIA, P. R.; DO NASCIMENTO, M. Z. Computational analysis of histological images from hematoxylin and eosin-stained oral epithelial dysplasia tissue sections. Expert Systems with Applications, Elsevier Ltd, v. 193, 5 2022. ISSN 09574174.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556, 2014.

SIQUEIRA, F. R.; SCHWARTZ W. R. ANS PEDRINI, H. Multi-scale gray level co-occurrence matrices for texture description. Neurocomputing, Elsevier, v. 120, p. 336–345, 2013.

SIRINUKUNWATTANA, K.; PLUIM, J. P. W.; CHEN, H.; X., Q. Gland segmentation in colon histology images: The glas challenge contest. MedicalImageAnalysis, v. 35, p. 489–502, 2017.

SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. Information Processing and Management, Elsevier, v. 45, p. 427–437, 2009.

SPANHOL, F. A.; CAVALIN, P. R.; OLIVEIRA, L. S.; PETITJEAN, C.; HEUTTE, L. Deep features for breast cancer histopathological image classification. 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, p. 1868–1873, 2017.

STRELNICKER, Y. M.; HAVLIN, S.; BUNDE, A. Fractals and percolation. Encyclopedia of Complexity and Systems Science, Springer New York, p. 3847–3858, 2009.

SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCKE, V.; RABINOVICH, A. Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.: s.n.], 2015. p. 1–9.

SZEGEDY, C.; VANHOUCKE, V.; IOFFE, S.; SHLENS, J.; WOJNA, Z. Rethinking the inception architecture for computer vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 12 2015.

TAINO, D. T.; RIBEIRO, M. G.; ROBERTO, G. F.; ZAFALON, G. F. D.; DO NASCIMENTO, M. Z.; TOSTA, T. A. A.; MARTINS, A. S.; NEVES, L. A. Analysis of cancer in histological images: employing an approach based on genetic algorithm. Pattern Analysis and Applications, Springer Science and Business Media Deutschland GmbH, v. 24, p. 483–496, 5 2021. ISSN 1433755X.

TALO, M. Convolutional neural networks for multi-class histopathology image classification. ArXiv, abs/1903.10035, 2019.

TAN, M.; LE, Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning, PMLR, v. 97, p. 6105–6114, 2019.

TANG, J.; RANGAYYAN, R. M.; XU, J.; NAQA, I. E.; YANG, Y. Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances. IEEE Transactions on Information Technology in Biomedicine, IEEE, v. 13, n. 2, p. 236–251, 2009.

TENGUAM, J. J.; LONGO, L. H. D. C.; SILVA, A. B.; FARIA, P. R.; NASCIMENTO, M. Z.; NEVES, L. A. Classification of H&E images exploring ensemble learning with two-stage feature selection. CFP2255E-ART, p. 1–4, 2022.

TENGUAM, J. J.; ROZENDO, G. B.; ROBERTO, G. F.; NASCIMENTO, M. Z.; MARTINS, A. S.; NEVES, L. A. Multidimensional and multiscale higuchi dimension for the analysis of colorectal histological images. 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, v. 1, p. 2833–2839, 2020.

TESAR, L.; SHIMIZU, A.; SMUTEK, D.; KOBATAKE, H.; NAWANO, S. Medical image analysis of 3d ct images based on extension of haralick texture features. Comput Med Imaging Graph., v. 32, n. 6, p. 513–520, 2008.

TORREY, L.; SHAVLIK, J. Transfer learning. In: SORIA, E.; MARTIN, J.; MAGDALENA, R.; MARTINEZ, M.; SERRANO, A. (Ed.). Handbook of Research on Machine Learning Applications. [S.l.: s.n.], 2009.

TRIPATHI, S.; SINGH, S. H. Ensembling handcrafted features with deep features: an analytical study for classification of routine colon cancer histopathological nuclei images. Multimedia Tools and Applications, Springer, 2020.

URBANOWICZ, R. J.; MEEKER, M.; La Cava, W.; OLSON, R. S.; MOORE, J. H. Relief-based feature selection: Introduction and review. Journal of Biomedical Informatics, v. 85, p. 189–203, 2018.

VARLEY, T. F.; CRAIG, M.; ADAPA, R.; FINOIA, P.; WILLIAMS, G.; ALLANSON, J.; PICKARD, J.; MENON, D. K.; STAMATAKIS, E. A. Fractal dimension of cortical functional connectivity networks & severity of disorders of consciousness. PLoS One, v. 15, n. 2, p. e0223812, 2020.

WAH, B. Wiley Encyclopedia of Computer Science and Engineering. [S.l.]: John Wiley & Sons, 2008. v. 5.

WANG, C.; YU, C. Automated morphological classification of lung cancer subtypes using h&e tissue images. Machine Vision and Applications, Springer, v. 24, p. 1383–1391, 2013.

WANG, Y.; ACS, B.; ROBERTSON, S.; LIU, B.; SOLORZANO, L.; WÄHLBY, C.; HARTMAN, J.; RANTALAINEN, M. Improved breast cancer histological grading using deep learning. Annals of Oncology, Elsevier Ltd, v. 33, n. 1, p. 89–98, jan 2022. ISSN 15698041.

WATANABE, K.; KOBAYASHI, T.; WADA, T. Semi-supervised feature transformation for tissue image classification. PLoS ONE, v. 11, n. 12, p. 1–20, 2016. ISSN 19326203.

WIBMER, A.; HRICAK, H.; GONDO, T.; MATSUMOTO, K.; VEERARAGHAVAN, H.; FEHR, D.; ZHENG, J.; GOLDMAN, D.; MOSKOWITZ, C.; FINE, S. W.; REUTER, V. E.; EASTHAM, J.; SALA, E.; VARGAS, H. A. Haralick texture analysis of prostate mri: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different gleason scores. European Radiology, Elsevier, v. 25, p. 2840–2850, 2015.

WILL, C. Does pooling in convolutional networks actually work? 2018. <https://principlesofdeeplearning.com/index.php/2018/08/27/is-pooling-dead-in-convolutional-nets-works/>. Online; acesso: 04-10-2020.

WOLFRAM. Wolfram Neural Net Repository. c2022. <https://resources.wolframcloud.com/NeuralNetRepository/>. Online; acesso: 18-08-2022.

XIA, R.; ZONG, C.; LI, S. Ensemble of feature sets and classification algorithms for sentiment classification. Information Sciences, v. 181, n. 6, p. 1138 – 1152, 2011.

XUE, B.; ZHANG, M.; BROWNE, W. N. Particle swarm optimisation for feature selection in classification: novel initialisation and updating mechanisms. Applied Soft Computing, Elsevier, v. 18, p. 261–276, 2014.

YAMAGUCHI, T.; HASHIMOTO, S. Fast crack detection method for large-size concrete surface images using percolation-based image processing. Machine Vision and Applications, v. 21, p. 797–809, 2010.

YU, C.; CHEN, H.; LI, Y.; PENG, Y.; LI, J.; YANG, F. Breast cancer classification in pathological images based on hybrid features. Multimed Tools Appl, v. 78, p. 21325–21345, 2019.

ZHANG, X.; ZHANG, Q.; CHEN, M.; SUN, Y.; QIN, X.; LI, H. A two-stage feature selection and intelligent fault diagnosis method for rotating machinery using hybrid filter and wrapper method. Neurocomputing, Elsevier, v. 275, p. 2426–2439, 2018. ISSN 0925-2312.

ZHAO, X.; LI, D.; YANG, B.; CHEN, H.; YANG, X.; YU, C.; LIU, S. A two-stage feature selection method with its application. Computers and Electrical Engineering, Elsevier Ltd., v. 47, p. 114–125, 2015. ISSN 00457906. Disponível em: <http://dx.doi.org/10.1016/j.compeleceng.2015.08.011>.

ZHENG, Y.; YANG, C.; MERKULOV, A. Breast cancer screening using convolutional neural network and follow-up digital mammography. In: MAHALANOBIS, A.; ASHOK, A.; TIAN, L.; PETRUCCELLI, J. C. (Ed.). Computational Imaging III. SPIE, 2018. v. 10669, p. 1 – 13. Disponível em: <https://doi.org/10.1117/12.2304564>.

Apêndice A

Resultados Obtidos

Tabela A.1: Acurácias e outras informações complementares obtidas nos 10 testes executados para cada combinação *ensemble* de descritores + seletor *wrapper* no *dataset* CR.

Seletor	PSO			GA			bGWO		
Execução	HF	DL	HF+DL	HF	DL	HF+DL	HF	DL	HF+DL
1	90,91%	98,79%	99,39%	93,94%	100%	98,79%	92,73%	100%	99,39%
2	92,73%	100%	99,39%	90,91%	98,18%	98,79%	94,55%	100%	99,39%
3	93,94%	100%	98,79%	90,91%	98,79%	99,39%	90,91%	100%	99,39%
4	90,91%	99,39%	99,39%	90,91%	99,39%	100%	92,12%	99,39%	99,39%
5	93,94%	98,79%	100%	90,30%	98,18%	100%	91,52%	100%	98,79%
6	91,52%	98,79%	100%	93,33%	98,79%	98,79%	92,12%	99,39%	99,39%
7	92,73%	100%	99,39%	93,33%	98,79%	99,39%	93,33%	100%	100%
8	90,30%	99,39%	100%	88,48%	100%	100%	89,70%	99,39%	100%
9	92,12%	100%	100%	89,70%	100%	100%	92,73%	99,39%	99,39%
10	90,30%	99,39%	99,39%	92,12%	99,39%	99,39%	93,33%	100%	99,39%
Média	91,94%	99,45%	99,58%	91,39%	99,15%	99,45%	92,30%	99,76%	99,45%
Desvio Padrão	1,30%	0,50%	0,39%	1,67%	0,67%	0,50%	1,30%	0,30%	0,33%
Melhor acurácia	93,94%	100%	100%	93,94%	100%	100%	94,55%	100%	100%

Tabela A.2: Acurácias e outras informações complementares obtidas nos 10 testes executados para cada combinação *ensemble* de descritores + seletor *wrapper* no *dataset* OED.

Seletor	PSO			GA			bGWO		
Execução	HF	DL	HF+DL	HF	DL	HF+DL	HF	DL	HF+DL
1	100%	97,97%	96,62%	98,65%	96,62%	96,62%	99,32%	96,62%	99,32%
2	100%	96,62%	99,32%	99,32%	97,97%	99,32%	99,32%	98,65%	99,32%
3	100%	95,95%	98,65%	100%	98,65%	99,32%	100%	97,97%	97,97%
4	100%	96,62%	99,32%	99,32%	97,97%	99,32%	100%	97,97%	99,32%
5	100%	97,30%	98,65%	99,32%	97,30%	100%	100%	96,62%	99,32%
6	100%	98,65%	97,97%	100%	96,62%	96,62%	99,32%	94,59%	98,65%
7	100%	97,97%	97,97%	100%	98,65%	97,30%	98,65%	94,59%	98,65%
8	100%	96,62%	99,32%	100%	97,97%	97,97%	100%	97,30%	99,32%
9	100%	96,62%	99,32%	100%	97,30%	98,65%	100%	99,32%	99,32%
10	100%	94,59%	97,30%	100%	97,30%	99,32%	100%	95,95%	99,32%
Média	100%	96,89%	98,45%	99,66%	97,64%	98,45%	99,66%	96,96%	99,05%
Desvio Padrão	0%	1,10%	0,91%	0,45%	0,69%	1,17%	0,45%	1,52%	0,45%
Melhor acurácia	100%	98,65%	99,32%	100%	98,65%	100%	100%	99,32%	99,32%

Tabela A.3: Acurácias e outras informações complementares obtidas nos 10 testes executados para cada combinação *ensemble* de descritores + seletor *wrapper* no *dataset* LA.

Seletor	PSO			GA			bGWO		
Execução	HF	DL	HF+DL	HF	DL	HF+DL	HF	DL	HF+DL
1	98,11%	93,56%	95,83%	98,67%	93,18%	95,64%	98,48%	93,37%	97,16%
2	98,67%	93,56%	97,73%	99,05%	93,56%	97,16%	99,24%	92,42%	96,59%
3	97,35%	93,75%	97,35%	98,48%	92,42%	96,59%	99,24%	95,45%	97,16%
4	98,30%	93,94%	96,59%	97,73%	92,80%	97,35%	98,48%	93,18%	96,97%
5	98,48%	93,94%	96,40%	98,30%	93,18%	97,35%	98,48%	93,18%	96,02%
6	98,86%	92,05%	97,16%	98,48%	92,80%	95,27%	98,86%	92,80%	95,27%
7	98,48%	93,37%	95,64%	98,67%	92,61%	97,35%	99,05%	94,13%	97,73%
8	98,48%	93,75%	96,78%	99,24%	92,99%	96,97%	98,48%	92,99%	95,08%
9	98,67%	93,56%	97,35%	98,67%	92,99%	96,40%	98,11%	94,13%	95,83%
10	97,54%	93,56%	97,73%	98,48%	93,56%	97,16%	98,86%	92,23%	96,59%
Média	98,30%	93,50%	96,86%	98,58%	93,01%	96,72%	98,73%	93,39%	96,44%
Desvio Padrão	0,47%	0,52%	0,70%	0,39%	0,35%	0,71%	0,36%	0,91%	0,82%
Melhor acurácia	98,86%	93,94%	97,73%	99,24%	93,56%	97,35%	99,24%	95,45%	97,73%

Tabela A.4: Acurácias e outras informações complementares obtidas nos 10 testes executados para cada combinação *ensemble* de descritores + seletor *wrapper* no *dataset* LG.

Seletor	PSO			GA			bGWO		
Execução	HF	DL	HF+DL	HF	DL	HF+DL	HF	DL	HF+DL
1	99,62%	98,49%	98,87%	99,25%	98,87%	98,87%	99,62%	98,49%	99,62%
2	99,25%	98,49%	98,87%	98,87%	98,87%	98,87%	98,87%	98,49%	99,62%
3	99,25%	98,11%	99,62%	99,62%	98,87%	98,11%	99,62%	98,87%	100%
4	99,62%	98,49%	98,87%	99,62%	97,74%	100%	99,25%	98,49%	99,25%
5	99,25%	98,87%	100%	99,62%	98,49%	98,87%	99,25%	98,49%	99,25%
6	99,62%	99,25%	99,62%	98,87%	98,87%	100%	99,62%	98,87%	99,62%
7	99,25%	99,25%	100%	99,25%	98,11%	99,62%	99,25%	98,87%	98,87%
8	98,87%	98,87%	98,87%	99,62%	99,62%	100%	99,25%	99,25%	99,62%
9	98,87%	99,25%	99,62%	99,25%	98,87%	99,25%	100%	99,25%	99,25%
10	100%	98,49%	99,62%	99,62%	99,25%	100%	98,87%	99,62%	99,62%
Média	99,36%	98,75%	99,40%	99,36%	98,45%	99,36%	99,36%	98,87%	99,47%
Desvio Padrão	0,34%	0,38%	0,45%	0,29%	0,51%	0,63%	0,34%	0,38%	0,30%
Melhor acurácia	100%	99,25%	100%	99,62%	99,62%	100%	100%	99,62%	100%

Tabela A.5: Acurácias e outras informações complementares obtidas nos 10 testes executados para cada combinação *ensemble* de descritores + seletor *wrapper* no *dataset* NHL.

Seletor	PSO			GA			bGWO		
Execução	HF	DL	HF+DL	HF	DL	HF+DL	HF	DL	HF+DL
1	78,61%	91,71%	91,98%	79,14%	91,44%	89,57%	79,41%	87,97%	90,11%
2	76,47%	88,24%	90,91%	76,47%	88,50%	89,57%	78,61%	92,25%	89,30%
3	78,88%	90,37%	89,30%	76,74%	90,64%	92,25%	78,61%	90,37%	91,44%
4	77,54%	89,57%	91,18%	78,34%	94,12%	91,18%	76,74%	88,77%	92,25%
5	77,01%	91,18%	89,30%	79,95%	89,84%	89,84%	77,54%	90,11%	90,11%
6	78,07%	91,44%	89,57%	77,54%	90,91%	89,57%	76,47%	91,44%	89,84%
7	79,14%	89,04%	92,25%	76,47%	90,37%	91,98%	78,61%	90,64%	90,64%
8	78,34%	90,64%	91,44%	77,81%	90,64%	90,11%	76,47%	89,57%	91,18%
9	78,34%	91,18%	89,04%	75,94%	88,77%	91,71%	80,21%	89,84%	90,64%
10	77,81%	93,05%	90,11%	77,54%	90,37%	91,44%	76,74%	90,37%	88,77%
Média	78,02%	90,64%	90,51%	77,59%	90,56%	90,72%	77,94%	90,13%	90,43%
Desvio Padrão	0,79%	1,33%	1,13%	1,21%	1,47%	1,04%	1,27%	1,16%	0,98%
Melhor acurácia	79,14%	93,05%	92,25%	79,95%	94,12%	92,25%	80,21%	92,25%	92,25%