

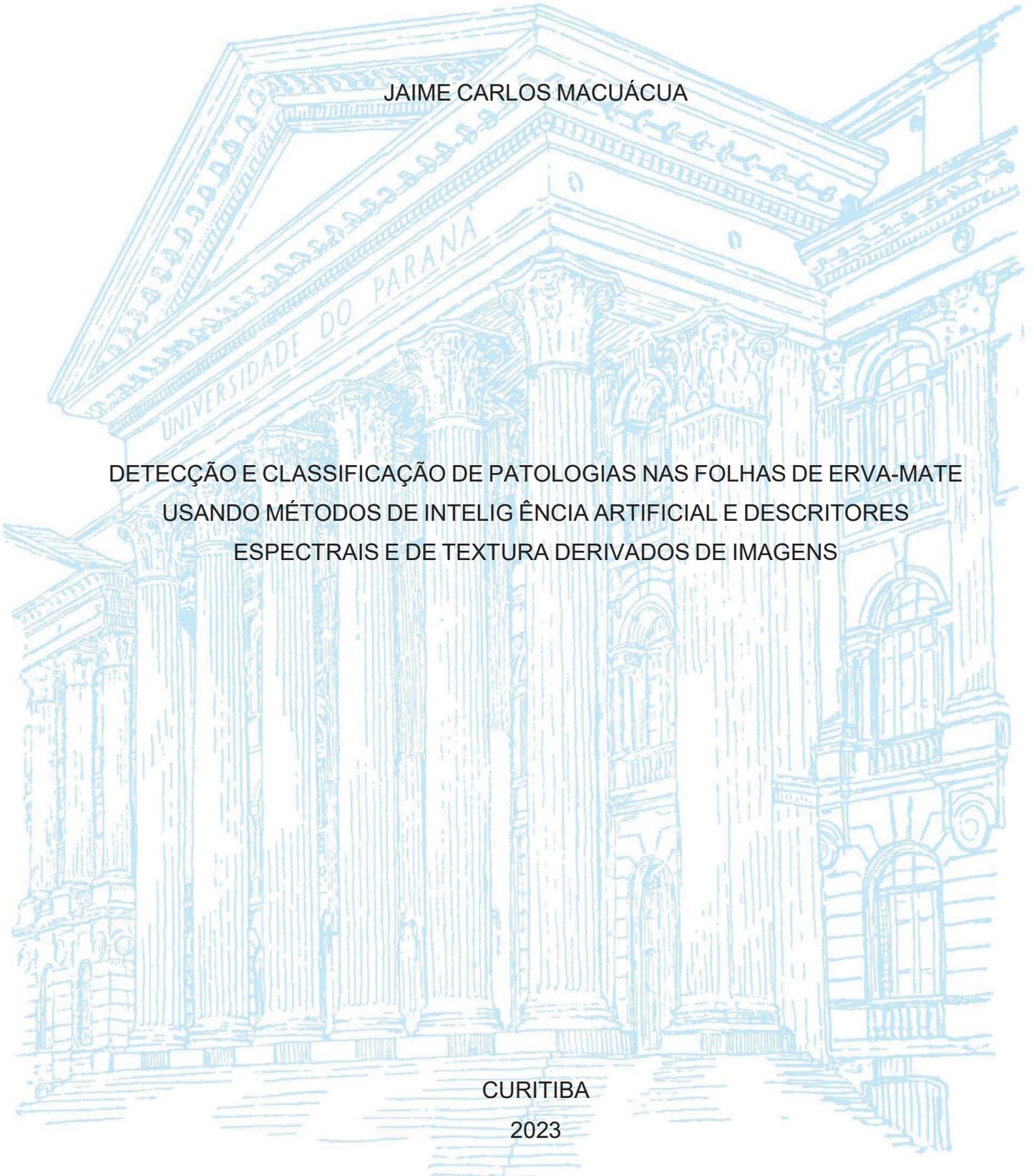
UNIVERSIDADE FEDERAL DO PARANÁ

JAIME CARLOS MACUÁCUA

DETECÇÃO E CLASSIFICAÇÃO DE PATOLOGIAS NAS FOLHAS DE ERVA-MATE
USANDO MÉTODOS DE INTELIGÊNCIA ARTIFICIAL E DESCRITORES
ESPECTRAIS E DE TEXTURA DERIVADOS DE IMAGENS

CURITIBA

2023



JAIME CARLOS MACUÁCUA

DETECÇÃO E CLASSIFICAÇÃO DE PATOLOGIAS NAS FOLHAS DE ERVA-MATE
USANDO MÉTODOS DE INTELIGÊNCIA ARTIFICIAL E DESCRITORES
ESPECTRAIS E DE TEXTURA DERIVADOS DE IMAGENS

Tese apresentada ao Programa de Pós-Graduação em Ciências Geodésicas, Setor de Ciências da Terra da Universidade Federal do Paraná, como requisito à obtenção do título de Doutor em Ciências Geodésicas.

Orientador: Prof. Dr. Jorge António Silva Centeno

CURITIBA

2023

FICHA CATALOGRÁFICA

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

Macuácuca, Jaime Carlos

Detecção e classificação de patologias nas folhas de erva-mate usando métodos de inteligência artificial e descritores espectrais e de textura derivados de imagens / Jaime Carlos Macuácuca. – Curitiba, 2023.

1 recurso on-line: PDF.

Tese (Doutorado) – Universidade Federal do Paraná, Setor de Ciências da Terra, Programa de Pós-Graduação em Ciências Geodésicas.

Orientador: Prof. Dr. Jorge Antônio Silva Centeno

1. Erva-mate – Doenças e pragas. 2. Rede neural convolucional.
3. Inteligência artificial. I. Centeno, Jorge Antônio Silva. II. Universidade Federal do Paraná. III. Programa de Pós-Graduação em Ciências Geodésicas. IV. Título.



TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação CIÊNCIAS GEODÉSICAS da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **JAIME CARLOS MACUÁCUA** intitulada: **DETECÇÃO E CLASSIFICAÇÃO DE PATOLOGIAS NAS FOLHAS DE ERVA-MATE USANDO MÉTODOS DE INTELIGÊNCIA ARTIFICIAL E DESCRITORES ESPECTRAIS E DE TEXTURA DERIVADOS DE IMAGENS**, sob orientação do Prof. Dr. JORGE ANTONIO SILVA CENTENO, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 20 de Novembro de 2023.

Assinatura Eletrônica

21/11/2023 16:24:54.0

JORGE ANTONIO SILVA CENTENO
Presidente da Banca Examinadora

Assinatura Eletrônica

20/11/2023 20:13:13.0

GABRIEL HENRIQUE DE ALMEIDA PEREIRA
Avaliador Externo (SIMEPAR)

Assinatura Eletrônica

23/11/2023 08:22:51.0

HIDEO ARAKI

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

20/11/2023 19:56:55.0

ANA PAULA DALLA CORTE

Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ)

DEDICATÓRIA

Dedico este trabalho a minha esposa e aos meus estimados filhos.

AGRADECIMENTOS

A Deus, que sempre iluminou a minha caminhada.

Aos meus pais, por todo o amor, carinho e educação ao longo dos anos, e por não me deixarem desistir.

Ao meu orientador Professor Jorge Antonio Silva Centeno pela indispensável ajuda e apoio, e por despertar meu interesse por esse projeto fascinante. O seu empenho foi essencial para a minha motivação à medida que as dificuldades iam surgindo ao longo do percurso.

Expresso minha gratidão a todos os profissionais do Departamento de Geomática do curso de Ciências Geodésicas da Universidade Federal de Paraná por todo o apoio que me deram ao longo da realização do meu trabalho.

Agradeço ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e à TWAS (The World Academy of Sciences) pela concessão da bolsa de estudos [Processo: 167085/2018-2].

À Universidade Eduardo Mondlane, Faculdade de Agronomia e Engenharia Florestal por todo apoio prestado.

Aos Colegas de curso pelo incentivo e troca de experiências.

À banca examinadora pela aceitação do convite e pelo tempo investido para leitura e avaliação desse projeto.

EPÍGRAFE

"Eu imagino um mundo onde a inteligência artificial nos permitirá ser mais produtivos, viver mais, e ter energia mais limpa" (Fei-Fei Li, 1999).

RESUMO

Nos últimos anos, a detecção e diagnóstico de doenças em plantas tornou-se uma grande preocupação na agricultura, impulsionando o interesse por soluções baseadas em inteligência artificial. Algoritmos de aprendizado profundo, destacaram-se na detecção de doenças a partir de imagens foliares, superando abordagens tradicionais. No entanto, a maioria das pesquisas concentram-se na classificação binária de imagens com manchas patológicas visíveis, resultando em modelos com baixa precisão e propensos a superajuste, devido à ênfase exclusiva em características texturais. Além disso, o uso de bancos de dados em condições reais de campo apresenta desafios, associados a fatores como, variações de iluminação, sombras, sobreposição entre folhas, oclusão e fundos complexos. Esta pesquisa visa desenvolver um sistema de visão computacional para detecção e diagnóstico automático de patologias em folhas de erva-mate, incluindo mapeamento espacial. A abordagem integra técnicas de segmentação, descritores de extração de características (cor, textura e forma) e aplicação da rede neural profunda do tipo CNN. Os modelos foram treinados com um banco de dados específico, contendo 4.660 imagens fotografadas em condições reais de campo, abrangendo 4 classes. A abordagem resultou numa precisão satisfatória em comparação com o uso de apenas um descritor, alcançando uma precisão global de 97,8% e uma acurácia de 97,6%, superando algoritmos de aprendizado de máquina, incluindo o classificador RF, que obteve acurácia de 86,4%. Os resultados são promissores e com potencial de serem ampliados no desenvolvimento de um sistema robusto de detecção e diagnóstico de patologias em plantas, apto em operar em ambientes reais de cultivo para suporte aos produtores rurais.

Palavras-chave: Doenças foliares de erva-mate; Imagens RGB; Segmentação; Descritores de cor, textura e forma; Rede Neural Convolutiva.

ABSTRACT

In recent years, the detection and diagnosis of plant diseases has become a major concern in agriculture, driving interest in solutions based on artificial intelligence. Deep learning algorithms have excelled in detecting diseases from leaf images, surpassing traditional approaches. However, research has focused on binary classification of images with visible pathological spots, resulting in models with low accuracy and prone to overfitting, due to the exclusive emphasis on textural features. In addition, the utilization of databases in real-field conditions presents challenges associated with factors such as lighting variations, shadows, leaf overlap, occlusion, and complex backgrounds. The study aims to develop a computer vision system for automatic detection and diagnosis of pathologies in yerba mate leaves, including spatial mapping. The approach integrates segmentation techniques, feature extraction descriptors (color, texture, and shape), and the application of a deep neural network, specifically a CNN (Convolutional Neural Network). The models were trained using a specific database, consisting of 4.660 images captured under real-field conditions, covering 4 classes. The approach resulted in satisfactory accuracy compared to the use of a single descriptor alone, achieving an overall accuracy of 97.8% and a precision of 97.6%, outperforming machine learning algorithms, including the RF classifier, which achieved an accuracy of 86.4%. The results are promising and have the potential to be expanded in the development of a robust system for detecting and diagnosing plant pathologies, capable of operating real growing environments to support rural producers.

Keywords: Mate leaf diseases; RGB images; Segmentation; Color, texture, and shape descriptors; Convolutional Neural Network.

LISTA DE FIGURAS

FIGURA 1 - ETAPAS DE SISTEMA DE PROCESSAMENTO DE IMAGENS	17
FIGURA 2 - REPRESENTAÇÃO ESQUEMÁTICA DO ESPAÇO RGB E HSV.....	26
FIGURA 3 - DIAGRAMA DE VENN DE ALGORITMOS DE APRENDIZADO DE MÁQUINA	30
FIGURA 4 - ESTRUTURA DE UMA CNN	38
FIGURA 5 - LOCALIZAÇÃO DA ÁREA DE ESTUDO	50
FIGURA 6 - FLUXOGRAMA METODOLÓGICO DA PESQUISA	52
FIGURA 7 - EXEMPLO DA PLANTA DE ERVA-MATE.....	53
FIGURA 8 - DIFERENTES CATEGORIAS DO CONJUNTO DE DADOS	56
FIGURA 9 - FIGURA 9 - COORDENADAS DE CONTORNOS DA FOLHA E CAIXA DELIMITADORA	57
FIGURA 10 - ESTRUTURA DA ARQUITETURA DA REDE YOLOV8	59
FIGURA 11 - REPRESENTAÇÃO DA CÉLULA DA GRADE QUE CONTÉM O OBJETO....	63
FIGURA 12 - ARQUITETURA DO MODELO U-NET	67
FIGURA 13 - ESTRUTURA MASK R-CNN PARA SEGMENTAÇÃO DE INSTÂNCIAS	68
FIGURA 14 - FLUXOGRAMA METODOLÓGICO PROPOSTO	71
FIGURA 15 - VISÃO GERAL DA ARQUITETURA CNN	78
FIGURA 16 - DIFERENTES ESPAÇOS DE CARACTERÍSTICAS DO MODELO SVM	81
FIGURA 17 -EXEMPLO DE METADADOS DE UMA IMAGEM.....	85
FIGURA 18 - RESULTADO DE DETECÇÃO E SEGMENTAÇÃO DE INSTÂNCIAS	86
FIGURA 19 - VISUALIZAÇÃO DE TENDÊNCIA DE PERDA E DO DESEMPENHO DO MODELO PARA DETECÇÃO DA FOLHA	88
FIGURA 20 - RESULTADOS DA SEGMENTAÇÃO DO FUNDO COMPLEXO DAS FOLHAS	89
FIGURA 21 - RESULTADO DE PRÉ-PROCESSAMENTO E SEGMENTAÇÃO DE MANCHAS VISÍVEIS DE DOENÇAS	93
FIGURA 22 - PRECISÃO E IMPRECISÃO DAS PREVISÕES DE DOENÇAS FOLIARES..	98
FIGURA 23 - INDICADORES DE ACURÁCIA NO TREINAMENTO E VALIDAÇÃO DO MODELO PROPOSTO	99
FIGURA 24 - INDICADORES DE PONTUAÇÃO DE PERDA NO TREINAMENTO E VALIDAÇÃO DO MODELO PROPOSTO.....	100
FIGURA 25 - ANÁLISE CURVA ROC-AUC DOS MODELOS	105
FIGURA 26 - DESEMPENHO DOS MODELOS NOS DIFERENTES EXPERIMENTOS....	108
FIGURA 27 - MAPA OCORRÊNCIA DE PATOLOGIAS NA ÁREA EXPERIMENTAL DE ERVA-MATE	112

LISTA DE TABELAS

TABELA 1- CARACTERÍSTICAS DA METODOLOGIA CIENTÍFICA APLICADA	9
TABELA 2 - PRODUÇÃO DE FOLHA VERDE DE ERVA-MATE, 2019-2021	13
TABELA 3 - MÉTODOS DE EXTRAÇÃO DE RECURSOS DE TEXTURA	20
TABELA 4 - DESCRIÇÃO DE DIFERENTES ESTUDOS COMPARATIVOS AO MODELO PROPOSTO.....	42
TABELA 5 - AVALIAÇÃO DE SEGMENTAÇÃO DAS FOLHAS DE ERVA-MATE.....	90
TABELA 6 - DELINEAMENTO DE EXPERIMENTOS COM BASE EM DESCRITORES.....	94
TABELA 7 - DESEMPENHO DO MODELO PROPOSTO COM OUTROS CLASSIFICADORES	95
TABELA 8 - MATRIZ DE CONFUSÃO DO MODELO HIBRIDO PROPOSTO.....	96
TABELA 9 - MATRIZ DE CONFUSÃO DE CLASSIFICADOR RANDOM FOREST COMBINANDO OS DESCRITORES.....	101
TABELA 10 - MATRIZ DE CONFUSÃO DE CLASSIFICADOR SUPPORT VECTOR MACHINE COMBINANDO OS DESCRITORES.....	103
TABELA 11 - MATRIZ DE CONFUSÃO DE CLASSIFICADOR K-NEAREST NEIGHBORS COMBINANDO OS DESCRITORES.....	104
TABELA 12 - ANÁLISE QUALITATIVA DOS MODELOS DE APRENDIZADO DE MÁQUINA NO USO ISOLADO E COMBINAÇÃO DOS DESCRITORES.....	106
TABELA 13 - ANÁLISE DE CURVA ROC-AUC DOS MODELOS EM VALIDAÇÃO CRUZADA.....	109
TABELA 14 - TEMPO DE TREINAMENTO COMBINANDO OS TRÊS DESCRITORES ...	110
TABELA 15 - DESCRIÇÃO DE DIFERENTES ESTUDOS COMPARATIVOS AO MODELO PROPOSTO.....	114
TABELA 16 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (A) - CLASSIFICADOR RF	144
TABELA 17 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (A) - CLASSIFICADOR SVM	144
TABELA 18 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (A) - CLASSIFICADOR KNN	144
TABELA 19 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (B) - CLASSIFICADOR RF	145
TABELA 20 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (B) - CLASSIFICADOR SVM	145

TABELA 21 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (B) - CLASSIFICADOR KNN	145
TABELA 22 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (C) - CLASSIFICADOR RF	146
TABELA 23 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (C) - CLASSIFICADOR SVM	146
TABELA 24 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (C) - CLASSIFICADOR KNN	146
TABELA 25 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (D) - CLASSIFICADOR RF	147
TABELA 26 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (D) - CLASSIFICADOR SVM	147
TABELA 27 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (D) - CLASSIFICADOR KNN	147
TABELA 28 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (E) - CLASSIFICADOR RF	148
TABELA 29 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (E) - CLASSIFICADOR SVM	148
TABELA 30 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (E) - CLASSIFICADOR KNN	148
TABELA 31 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (F) - CLASSIFICADOR RF	149
TABELA 32 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (F) - CLASSIFICADOR SVM	149
TABELA 33 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (F) - CLASSIFICADOR KNN	149
TABELA 34 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (G) - CLASSIFICADOR RF	150
TABELA 35 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (G) - CLASSIFICADOR SVM	150
TABELA 37 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (G) - CLASSIFICADOR KNN	150

LISTA DE QUADROS

QUADRO 1 - CARACTERIZAÇÃO DE DOENÇAS DE CULTURA DE ERVA-MATE	14
--	----

LISTA DE ABREVIATURAS E SIGLAS

Adam	<i>Adaptive Moment Estimation</i>
ANN	Rede Neural Artificial
AUC	Area Under the Curve
BGR	<i>Blue-Green-Red</i>
CASC-	<i>Comprehensive Automation for Specialty Crops - Internal Feeding</i>
IFW	<i>Worm Database</i>
CBS	<i>Conv-BN-SiLU</i>
CNN	<i>Convolutional Neural Network</i>
CVAT	<i>Computer Vision Annotation Tool</i>
DL	<i>Deep Learning</i>
DNA	Ácido Desoxirribonucleico
DNNs	<i>Deep Neural Networks</i>
DT	<i>Tree Dcision</i>
ELAN	<i>Extraordinary Linear Attention Network</i>
EXIF	<i>Exchangeable Image File Format</i>
FDA	<i>Fisher Discriminant Analysis</i>
FN	Falsos Negativos
FP	Falsos Positivos
FPN	<i>Feature Pyramid Networks</i>
GAN	<i>Generative adversarial networks</i>
GIS	<i>Geographic Information System</i>
GLCM	<i>Gray-Level Cooccurrence Matrix</i>
GNSS	<i>Global Navigation Satellite System</i>
GPS	<i>Global Positioning System</i>
GPU	<i>Graphics Processing Unit</i>
GrabCut	<i>Graph Cut</i>
HAAR	<i>Haar Wavelet</i>
HDCT	<i>High Dimensional Color Transform</i>
HOG	Histograma de um Gradiente Orientado

HSV	Hue Saturation Value
IA	Inteligência Artificial
IBGE	Instituto Brasileiro de Geografia e Estatística
ILSVRC	<i>ImageNet Large Scale Visual Recognition Challenge</i>
IoU	<i>Intersection Over Union</i>
IRA	Índice de Refletância de Antocianina
IRF	Índice de Refletância Fotoquímica
KNN	K-Nearest Neighbors
LBP	Padrão Binário Local
LDA	Linear Discriminant Analysis
mAP	precisão média
MCC	Matriz de Coocorrência de Cores
ML	<i>Machine Learning</i>
NB	Naive Bayes
NIR	<i>Near Infrared</i>
NMS	<i>Non-Maximum Suppression</i>
ORB	<i>Oriented FAST and Rotated BRIEF</i>
PAN-FPN	<i>Path Aggregation Network with Feature Pyramid Network</i>
PCA	Principal Component Analysis
PHOG	Pirâmide de Histograma de Gradiente de Orientação
PIB	Produto Interno Bruto
PVD	<i>PlantVillage</i>
QGIS	<i>Quantum Geographic Information System</i>
RAM	<i>Random Access Memory</i>
R-CNN	<i>Region based Convolutional Neural Network</i>
ReLU	<i>Rectified Linear Units</i>
RF	<i>Random Forest</i>
RGB	<i>Red, Green, Blue</i>
RL	Regressão Logística
RNA	Rede Neural Artificial
ROC	<i>Receiver Operating Characteristic</i>
ROI	Identificação de Região de Interesse

RoIPool	<i>Region of Interest Pooling</i>
RPN	Rede de Região Proposta
SGD	<i>Stochastic Gradient Descent</i>
SIFT	<i>Scale Invariant Feature Transform</i>
SIG	Sistema de Informação Geográfica
SiLu	<i>Sigmoid Linear Unit</i>
SPPF	<i>Spatial Pyramid Pooling Fusion</i>
SPPR	<i>Spatial Pyramid Pooling Fusion</i>
SSD	<i>Single Shot Detector</i>
SURF	<i>Speeded Up Robust Features</i>
SVM	<i>Support Vector Machines</i>
UCI	<i>University of California Irvine Machine Learning Repository</i>
U-net	<i>U-shaped architecture of the network</i>
VN	Verdadeiros negativos
VP	Verdadeiros positivos
YOLO	<i>You Only Look Once</i>

LISTA DE SÍMBOLOS

'	minutos
"	segundos
%	porcentagem
°	graus
GB	gigabyte
km	Quilômetro
m	metro
s	segundo

SUMÁRIO

1	INTRODUÇÃO E CARACTERIZAÇÃO DO PROBLEMA.....	1
1.1	HIPÓTESE	4
1.2	OBJETIVOS	5
1.2.1	Objetivo Geral	5
1.2.2	Objetivos Específicos.....	5
1.3	JUSTIFICATIVA	6
1.4	ESTRUTURA DA TESE	8
1.5	ABORDAGEM METODOLÓGICA.....	9
2	REVISÃO DA LITERATURA.....	10
2.1	FITO DIAGNOSE DIGITAL	10
2.2	ERVA-MATE	12
2.3	PROCESSAMENTO DIGITAL DE IMAGENS	16
2.3.1	Aquisição	17
2.3.2	Pré-Processamento	18
2.3.3	Segmentação.....	18
2.3.4	Representação e descrição	19
2.3.5	Classificação.....	23
2.4	SISTEMAS DE CORES.....	25
2.5	INTELIGÊNCIA ARTIFICIAL	27
2.5.1	Aprendizado de Máquina no Diagnóstico de Doenças de Plantas	31
2.5.2	Aprendizado Profundo e Técnicas de Transferência de Aprendizado	33
2.5.2.1	Detecção e segmentação	34
2.5.2.2	Classificação.....	37
2.6	MAPEAMENTO DE PATOLOGIAS DE PLANTAS USANDO IMAGENS DIGITAIS	48
3	MATERIAIS E MÉTODOS	50
3.1	CARACTERIZAÇÃO DA ÁREA DE ESTUDO.....	50
3.2	PROCEDIMENTO METODOLÓGICO.....	51
3.2.1	Coleta de Imagens e Composição do Banco de Dados.....	52

3.2.2 Detecção.....	57
3.2.3 Avaliação do desempenho dos modelos de detecção e segmentação	69
3.2.6 Aprendizado Profundo	76
3.2.7 Comparação com outros classificadores	78
3.2.8 Avaliação do desempenho dos modelos de classificação	82
3.3 MAPEAMENTO DA OCORRÊNCIA DE PATOLOGIAS.....	84
4 RESULTADOS.....	86
4.1 DETECÇÃO DA FOLHA DE ERVA-MATE.....	86
4.2. SEGMENTAÇÃO DO FUNDO COMPLEXO DA FOLHA DE ERVA-MATE.....	89
4.2.1 Comparação de diferentes métodos de segmentação.....	90
4.3 CLASSIFICAÇÃO DE PATOLOGIAS USANDO DESCRITORES DE EXTRAÇÃO DE CARACTERÍSTICAS E CNN.....	91
4.3.1 Resultados de pré-processamento e segmentação.....	91
4.3.2 Delineamento de experimentos com base em descritores	94
4.3.3 Resultados de classificação do modelo proposto e outros modelos de aprendizado de máquina	95
4.3.4 Avaliação qualitativa dos descritores.....	105
4.3.5 Tempo de processamento	109
4.4 MAPEAMENTO DA OCORRÊNCIA DE PATOLOGIAS NA ÁREA DE ESTUDO	111
5 DISCUSSÃO DOS RESULTADOS	113
5.1 DETECÇÃO E SEGMENTAÇÃO DA FOLHA DE ERVA-MATE.....	113
5.2 CLASSIFICAÇÃO DE PATOLOGIAS DE ERVA-MATE.....	115
6 CONSIDERAÇÕES FINAIS	120
6.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS	122
7 REFERÊNCIAS.....	123
APÊNDICES – MATRIZES DE CONFUSÃO DOS EXPERIMENTOS	144

1 INTRODUÇÃO E CARACTERIZAÇÃO DO PROBLEMA

Nos países em desenvolvimento, mais de 70% da produção agrícola é gerada por pequenos agricultores (Bhat, 2021; UNICEF, 2021). Esse cenário também se reflete na produção de alimentos consumidos no Brasil. De acordo com Dhaka et al. (2021), a produtividade e a qualidade de produção de plantas são influenciadas por vários fatores, alguns relacionados às condições climáticas, como temperatura, umidade e precipitação, e fatores associados ao surgimento de bactérias, vírus, fungos, nematóides e outros microrganismos. Considerando estes últimos fatores, estudo como o de Liakos (2018) documentam que existe uma perda significativa da produção e produtividade devido a doenças e pragas. As doenças e pragas afetam o crescimento normal das plantas, podendo chegar a causar a morte da planta (Liu e Wang, 2021).

A detecção precoce e eficaz de doenças e pragas é essencial para que os agricultores possam tomar as medidas necessárias para combater esses problemas ou minimizar seus impactos. Isso se torna ainda mais crucial, uma vez que tais problemas são responsáveis por causar uma redução significativa na produção agrícola (Dhaka et al., 2021). Essa diminuição na produção representa uma ameaça significativa para a segurança alimentar em nível global e afeta diretamente o sustento de famílias que dependem da agricultura como sua principal fonte de renda.

As doenças que afetam as plantas se manifestam por meio de diversos sintomas visíveis, tais como lesões, alterações na coloração, danos às folhas e ao caule, crescimento anormal do caule, folhas, botões, flores e/ou raízes. Além disso, as folhas podem exibir sinais como manchas, ressecamento e queda prematura, sendo indicadores claros de doença (Chouhan et al., 2020). A eficácia do controle fitossanitário de doenças em plantas está intrinsecamente ligada à rápida e precisa identificação do agente causador (Chen et al., 2020; Dhaka et al., 2021) e o mapeamento das ocorrências para prever sua distribuição espacial. A capacidade de prever com precisão a ocorrência dessas doenças em tempo hábil desempenha um papel crucial no suporte à implementação de medidas preventivas e protetivas

adequadas. Esse processo contribui significativamente para melhorar a qualidade do rendimento das culturas e aumentar a produtividade (Hu et al., 2019).

Várias estratégias têm sido desenvolvidas com o apoio de serviços de extensão agrícola com o objetivo de reduzir as perdas na produção. Uma das abordagens mais notáveis é a histórica utilização de pesticida, que, ao longo do tempo, tem sido complementada pelo conceito de Manejo Integrado de Pragas e Doenças. Essa abordagem tem consistentemente demonstrado sua relevância, destacando a importância da identificação e classificação precisa e precoce das doenças como um elemento essencial para a gestão eficiente das patologias que afetam as plantas e, por conseguinte, para a melhoria da produção agrícola, esforço também que é apoiado pelo avanço e disseminação de tecnologia neste setor.

Para uma identificação adequada desses problemas, a participação de agrônomos ou outros especialistas é fundamental. No entanto, em países em desenvolvimento, os pequenos agricultores nem sempre têm acesso a esse tipo de suporte. Portanto, a análise e o reconhecimento de sintomas de doenças de plantas por parte dos próprios agricultores representam um desafio considerável devido à falta de infraestruturas, como laboratórios especializados em diagnóstico de doenças em plantas, porém, isso implica em custos proibitivos para pequenos produtores. Além disso, essa abordagem exige um conhecimento técnico sólido no processo de diagnóstico.

De acordo com Bai et al. (2018) e Chen et al. (2020), a abordagem de diagnóstico de doenças baseada na inspeção visual pode levar a previsões incorretas devido a ilusões visuais e tendências subjetivas. Além disso, essa abordagem é geralmente considerada impraticável em grandes áreas de terras agrícolas. O estudo realizado pela Agência Brasil (2015)¹, mostrou que a precisão do método de inspeção visual de doenças em plantas é relativamente baixa, em torno de 50% em relação às plantas com sintomas, em contraste com a inspeção automatizada, que alcança uma

¹Agência Brasil (2015) - Protótipo criado pela Embrapa detecta doença que afeta pomares de laranja. <https://agenciabrasil.ebc.com.br/pesquisa-e-inovacao/noticia/2015-10/prototipo-criado-pela-embrapa-detecta-doenca-que-afeta-pomares-de-laranja#>

precisão de 80-90%. De acordo com Kittler et al. (2017), a precisão da inspeção visual não assistida é aproximadamente 60%, mas pode aumentar para 75%-84% com o auxílio de sistemas automatizados e profissionais treinados.

O desenvolvimento de ferramentas de processamento de imagens revolucionou a capacidade de analisar imagens para detectar objetos e extrair informações com base em sua forma ou cor. Um exemplo disso é a utilização dos descritores de extração de características como GLCM- Matrizes de Coocorrência de Níveis de Cinza (Saha and Ahsan, 2021; Vishnoi et al., 2022); LBP- Padrão Binário Local (Kaur, 2021; Mathew et al., 2021); HOG- Histograma de Gradientes Orientados (Pothen and Pai, 2020; Ochango et al., 2022,); SIFT- Transformação de Características Invariantes à Escala (Gayathri Devi e Neelamegam, 2019; Chouhan et al., 2021). Outras técnicas de extração de características são baseadas em sistemas de cores, com destaque a Histograma de Cores (Basavaiah and Anthony, 2020; Xian e Ngadiran, 2021); Momentos de Cor (Chaki and Dey, 2020; Anton et al., 2021) e MCC- Matriz de Coocorrência de Cores (Khitthuk et al., 2018; Anton et al., 2021), entre outros. Esse processo está sendo aprimorado com o surgimento de métodos de inteligência artificial. Atualmente, os métodos de aprendizado profundo (Deep Learning - DL), especialmente aqueles baseados em rede neural profunda do tipo CNN (Convolutional Neural Network), ganharam ampla aplicação no diagnóstico de doenças de plantas (Abdu et al., 2020; Sujatha et al., 2021; Li et al., 2021; Saberi Anari, 2022; Chen et al., 2022 e Singh et al., 2022). Os métodos de aprendizagem profundo são capazes de aprender representações hierárquicas de características em vários níveis de abstração, desde as bordas, texturas e padrões de imagens. Isso permite que eles descubram padrões complexos e nuances nos dados, tornando-os extremamente eficazes em tarefas como detecção e classificação - representam tecnologia avançada neste campo.

Os modelos de aprendizagem profunda podem ser usados no diagnóstico de doenças em condições reais de campo de cultivo. Em situações envolvendo cultivos em larga escala, é possível combinar esse sistema com veículos agrícolas autônomos, exemplo de VANTs (Veículos Aéreos Não Tripulados). Isso possibilita a

localização precisa e ágil de problemas fitopatológicos em toda a extensão do campo de cultivo, por meio da captura contínua de imagens (Johannes et al., 2017).

Esta pesquisa se insere na aplicação de métodos de inteligência artificial, especificamente da rede neural profunda do tipo CNN, para a detecção e classificação de patologias nas folhas de erva-mate. A erva-mate (*Ilex paraguariensis* A. St.-Hill) é uma espécie florestal de ocorrência natural explorada com finalidade não madeireira. Pertence à família Aquifoliaceae - ela é nativa da Região Sul do Brasil, bem como da Argentina e do Paraguai. A escolha dessa cultura se justifica pela sua importância na produção agrícola do Estado do Paraná, e o estímulo do fomento desta cultura atrelado à difusão tecnológica pode contribuir para redução de níveis de emissões de gases de efeito estufa. Os resultados desta pesquisa podem oferecer contribuições significativas para o desenvolvimento de sistema inteligente para o diagnóstico precoce de doenças, melhorando as condições de vida dos agricultores que carecem de assistência especializada.

1.1 HIPÓTESE

Dada a facilidade de obtenção de imagens de folhas por meio de dispositivos móveis, como *smartphones*, e considerando que os modelos de aprendizagem profunda têm a capacidade de identificar objetos ou partes deles em imagens, é possível adaptar esses métodos para a detecção e diagnóstico de doenças nas folhas de erva-mate. No entanto, considerando as dificuldades encontradas por esses métodos ao lidar com imagens de folhas de espécies específicas obtidas em ambientes reais de campo de cultivo, a hipótese é que ao incorporar parâmetros texturais e espectrais derivados das imagens, a acurácia dos métodos de diagnóstico pode ser aumentada.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

- Desenvolver uma metodologia para identificação e classificação automático, incluindo o mapeamento espacial de patologias nas folhas de cultura de erva-mate usando métodos de inteligência artificial e análise qualitativa de descritores de extração de características derivados das imagens fotografadas.

1.2.2 Objetivos Específicos

- Desenvolver um método para segmentar imagens de folhas de erva-mate fotografadas em condições reais de campo de cultivo;
- Avaliar o desempenho do método escolhido com outros modelos de segmentação de folhas de erva-mate em termos de acurácia;
- Avaliar o uso integrado de descritores espaciais, espectrais e textura na caracterização de patologias nas imagens de folhas de erva-mate;
- Construir modelos de detecção e diagnóstico de doenças de plantas baseados na integração dos descritores e algoritmos de aprendizagem de máquina e aprendizado profundo, para efeito de comparação;
- Mapear a distribuição espacial de ocorrência de patologias com base nas informações geoespaciais fornecidas pela câmera.

1.3 JUSTIFICATIVA

A agricultura desempenha um papel de extrema relevância na economia global, sendo a agricultura brasileira notoriamente reconhecida por sua alta competitividade e por sua capacidade de gerar empregos, riqueza, alimentos, fibras e bioenergia, tanto para o Brasil quanto para outras nações. Este setor representa uma das maiores contribuições para o crescimento do Produto Interno Bruto (PIB) nacional, representando aproximadamente 21% do total da produção econômica do país. Adicionalmente, a agricultura é responsável por um quinto de todos os empregos e contribui com 43,2% das exportações, atingindo um valor de US\$ 96,7 bilhões em 2019 (Brasil, 2020). Nas últimas 4 décadas, de 1975 a 2019, a produção de grãos experimentou um impressionante aumento, passando de 38,1 milhões de toneladas para 232,6 milhões de toneladas, representando um crescimento impressionante de 510% (IBGE, 2020). A produção de carnes também registrou um notável crescimento, passando de 2,9 milhões de toneladas para 27,9 milhões de toneladas, o que equivale a um aumento substancial de 858% no mesmo período (United States, 2020). Além disso, o setor florestal aumentou sua produtividade em mais de 150%, destacando-se o cultivo de espécies como Pinus e Eucalyptus. A cafeicultura, nos últimos 25 anos, testemunhou um aumento na produtividade de mais de quatro vezes (Embrapa, 2020: 31 p)².

Dentre as culturas que se destacam significativamente em termos de faturamento no Brasil, incluem-se o algodão, arroz, cacau, café, cana, feijão, milho, soja, trigo e madeira destinada à produção de celulose. Além dessas, existem outras culturas que apresentam forte competitividade econômica e mostram-se promissoras, muitas das quais são originárias das distintas regiões do Brasil e, portanto, contribuem para a sustentabilidade econômica. Um exemplo notável é a cultura de erva-mate (CEPEA/USP e CNA, 2020). A produção e consumo de erva-mate é tradicional na

² VII Plano Diretor da Embrapa : 2020–2030 / Embrapa. – Brasília, DF : Embrapa, 2020. 31 p. : il. color. ; ? cm x ? cm.

região sul do Brasil, abrangendo estados como Mato Grosso do Sul, Rio de Janeiro, Paraná, Santa Catarina e Rio Grande do Sul, e sua industrialização desempenha um papel econômico significativo no país (IBGE, 2021). No entanto, apesar da importância financeira dessa cultura para a economia local, as plantações enfrentam desafios decorrentes da redução na produção e do declínio das plantas, ocasionados por diversas doenças, incluindo cercosporiose, pinta-preta, fumagina, fuligem, podridão-do-tronco, nematoide e podridão das raízes, entre outras. Esses problemas têm causado prejuízos financeiros consideráveis na indústria agrícola da erva-mate.

O diagnóstico preciso de doenças em plantas é fundamental. Os métodos tradicionais apresentam limitações, como a inspeção visual e testes laboratoriais, o que os tornaram dispendiosos, demorados, exigirem conhecimentos especializados e serem suscetíveis a subjetividade. Conforme destacado por Ferentinos (2018), o diagnóstico de doenças em plantas com base na observação óptica dos sintomas nas folhas apresenta um alto grau de complexidade. Dada essa complexidade e a vasta extensão de cultivos e seus desafios fitopatológicos, até mesmo agrônomos e fitopatologistas experientes frequentemente têm dificuldade em identificar doenças específicas, levando a conclusões e tratamentos equivocados (Mohanty et al., 2016; Yang e Guo, 2017; Liakos, 2018).

Ou por outra, as técnicas convencionais apresentam baixa precisão devido à extração inadequada de características. Além disso, a variação da iluminação, o ruído e a qualidade da imagem são os principais fatores que criam complexidades adicionais durante o processo de treinamento. Com base nessas suposições, esta pesquisa se concentra no desenvolvimento de uma nova abordagem para detecção de doenças foliares de plantas, desenvolvendo um esquema combinado de segmentação, extração de características e classificação de doenças.

Nesse contexto, o desenvolvimento de metodologia para sistema automatizado para a detecção e classificação de doenças de plantas representaria uma ferramenta extremamente valiosa para os agrônomos encarregados de realizar esses diagnósticos por meio inspeção visual das folhas das plantas infetadas. Tal sistema ofereceria benefícios diretos tanto para os profissionais responsáveis pelo

diagnóstico quanto para os produtores locais. Busca-se, assim, uma solução tecnológica que permita a identificação rápida de doenças na cultura da erva-mate, visando aprimorar tanto a qualidade quanto a eficiência da produção.

1.4 ESTRUTURA DA TESE

Dada as contribuições descritas acima, a presente tese está estruturada da seguinte maneira:

1. **Introdução e caracterização do problema:** este item inicial tem como objetivo contextualizar o trabalho e o problema a ser estudado, apresentando as lacunas e ideias iniciais sobre a solução do mesmo. Além disso, são apresentados os objetivos da realização da Tese de Pós-graduação em Ciências Geodésicas;
2. **Revisão da literatura:** aqui foram apresentados os assuntos a serem tratados na pesquisa, tais como principais conceitos em torno de diagnose digital, processamento de imagem digital, inteligência artificial e algumas metodologias utilizadas anteriormente;
3. **Materiais e Métodos:** neste item são apresentados os métodos e o procedimento metodológico do sistema sugerido para a resolução do problema em estudo;
4. **Resultados:** neste item são apresentados os resultados obtidos nos experimentos;
5. **Discussão dos resultados:** aqui foram apresentadas as discussões, comparando com outros trabalhos da literatura;
6. **Considerações finais:** aqui foram apresentadas as conclusões relativas aos resultados.

7. **Recomendações:** neste item foram apresentadas recomendações para direcionar trabalhos futuros. Estas abrangem aspectos voltados para a melhoria do modelo, bem como exploram outras direções de pesquisa.

1.5 ABORDAGEM METODOLÓGICA

A pesquisa enquadra-se no campo de pesquisa Teórico-Aplicada Quantitativa. Em sentido amplo, a presente pesquisa foi caracterizada pela metodologia descrita na Tabela 1.

TABELA 1- CARACTERÍSTICAS DA METODOLOGIA CIENTÍFICA APLICADA

Classificação	Tipo
Quanto à natureza	Pesquisa aplicada
Quanto aos objetivos	Exploratórios e descritivos
Quanto às abordagens	Quantitativas
Quanto aos procedimentos	Experimentais

Fonte: Autor (2023).

2 REVISÃO DA LITERATURA

A presente pesquisa visa implementar técnicas de aprendizagem de máquina e aprendizagem profunda para a detecção e classificação de doenças de plantas em condições reais de campo de plantio. Antes de se aprofundar em detalhes específicos no estudo e com base na literatura existente, é fundamental compreender os principais conceitos relacionados à diagnose digital, processamento digital de imagens, inteligência artificial e algumas metodologias previamente empregadas, a fim de entender suas limitações e vantagens.

2.1 FITO DIAGNOSE DIGITAL

As doenças em plantas não são consequência de uma alteração isolada, é resultado da interação entre uma planta suscetível, um agente patogênico e o meio ambiente. A sintomatologia é a parte da fitopatologia que se dedica ao estudo dos sintomas e sinais que caracterizam uma determinada doença em plantas. Sintomas são as respostas da planta (hospedeiro) a qualquer manifestação de agentes nocivos. Estes são definidos como sinais, considerados estruturas do patógeno quando são visíveis no tecido doente (Carollo e Filho, 2016, pp.14).

De acordo com Carollo e Filho, (2016, pp.14-16), os sintomas mais comuns associados às doenças de plantas são: (i) *Anasarca* – caracterizado pelo extravasamento de conteúdo celular, resultando em áreas de aspecto encharcado; (ii) *Cancro* – manifesta-se por lesões necróticas que formam depressões nos tecidos corticais dos caules, tubérculos e raízes; (iii) *Clorose* – representa a ausência parcial ou total da coloração verde normal das folhas, podendo tornar-se verde-amarelada ou até mesmo esbranquiçada; (iv) *Galha* – refere-se ao desenvolvimento anormal de tecidos devido à hipertrofia, resultando em supercrescimento de células; (v) *Gomose* – envolve a exsudação de goma a partir de lesões, principalmente em caules ou frutos; (vi) *Mancha* – mais comuns em folhas, mas também podem ocorrer em flores, frutos

ou ramos – as manchas foliares têm formas variadas, podendo ser irregulares, angulares, circulares; *(vii) Mosaico* – caracterizado por áreas cloróticas intercaladas com áreas de verde mais escuro, geralmente observadas em folhas; *(viii) Murcha* – resulta na perda de turgescência de folhas, pecíolos e hastes suculentas, decorrente da obstrução do sistema vascular ou destruição do sistema radicular; *(ix) Necrose* – corresponde ao escurecimento de tecidos causado pela morte de células; *(x) Podridão* – envolve a morte e desintegração de tecidos, decorrente da atividade enzimática de fitopatógenos; *(xi) Pústula* – manifesta-se como pequenas manchas necróticas, geralmente menor do que 1,0 cm, com elevação da epiderme, que se rompem devido à produção e exposição de esporos fúngicos; *(xii) Requeima* – caracteriza-se pela necrose repentina de órgãos aéreos, como folhas, flores e brotações; *(xiii) Tombamento* – resulta da podridão dos tecidos tenros da base de seu caulículo e; *(xiv) Verrugose* – envolve o crescimento excessivo de tecidos epidérmicos e corticais, geralmente modificados pela ruptura e suberificação das paredes celulares, originando lesões salientes e ásperas, comumente encontrados em frutos, tubérculos e folhas.

A diagnose de doenças em plantas é uma tarefa complexa e abrangente que requer a coleta de diversas informações para apoiar a tomada de decisões relacionadas aos tratamentos culturais. Além disso, o sucesso no tratamento de uma doença depende principalmente da rapidez e eficiência do seu diagnóstico (Carollo e Filho, 2016). O diagnóstico de doenças em plantas acarreta várias consequências, que vão desde os impactos econômicos e sociais nas áreas de cultivo até efeitos que ultrapassam o âmbito agrônomo imediato, podendo afetar produtores, comunidades rurais, consumidores, o Estado e o meio ambiente (Carollo e Filho, 2016, pp.20).

Nesse contexto, a tecnologia desempenha um papel crucial, uma vez que, muitas vezes, a assistência de um profissional não está prontamente disponível. Assim, esforços têm sido direcionados para o desenvolvimento de ferramentas computacionais que possam auxiliar no combate aos problemas fitossanitários. O objetivo principal é reduzir o tempo necessário para implementar as primeiras medidas de controle, especialmente quando não se dispõe de um especialista. Essa abordagem é particularmente relevante no contexto da presente pesquisa, que visa o

diagnóstico de doenças como *cercosporiose*, *pinta-preta*, *antracnose*, *fuligem* - doenças com maior destaque na cultura da erva-mate.

2.2 ERVA-MATE

A erva-mate é conhecida por sua forte cultura de consumo em diversas formas, incluindo chás, chimarrão³, além de ser utilizada em produtos cosmético, de limpeza e higiene. O Brasil, apesar de possuir a maior área de ervais nativos, é o segundo maior produtor mundial da planta. O aumento na demanda pelo consumo de erva-mate levou à substituição de ervas nativas por plantações cultivadas, o que, por sua vez, aumentou a suscetibilidade da cultura a pragas e doenças (Piassetta et al., 2021). A mudança de sistema de produção de erva-mate trouxe desafios adicionais em termos de manejo fitossanitário. Isso ressalta a importância da pesquisa e diagnóstico eficazes de doenças que afetam essa cultura.

A Tabela 2 apresenta os dados de produção de folhas de erva-mate. A cadeia produtiva da erva-mate se estende por todo o território brasileiro, mas é na região sul do Brasil - abrangendo Mato Grosso do Sul, Rio de Janeiro, Paraná, Santa Catarina e Rio Grande do Sul - que ela se destaca, com uma longa tradição tanto na produção quanto no consumo de erva-mate. A industrialização da erva-mate nesses estados representa um fator econômico relevante para o país (IBGE, 2021; SIDRA/IBGE, 2022). No período de 2019, 2020 e 2021, esses estados tiveram área total de produção de erva-mate de 70.305ha, 72.122ha e 72.106ha, com produção de 522.259t, 527.546t, 557.987t, respectivamente. O consumo de chimarrão consolidou a extensa cadeia produtiva da erva-mate. De acordo com os dados do SIDRA/IBGE (2022), o valor de produção da safra de erva-mate para os estados mencionados no mesmo período foi de 792.71 milhões de reais em 2019, 884.05 milhões de reais em 2020 e 900.84 milhões de reais em 2021. Ainda os dados de SIDRA/IBGE (2022)

³ No dia 24 de abril é comemorado o dia do chimarrão (Lei Estadual 11.929/2003), principal hábito de consumo da erva-mate no sul do Brasil, sendo bebida oficial do RS.

indicam que o estado de Paraná registrou um notável crescimento de 60,48% na produção por hectare, e em relação ao faturamento, no mês de fevereiro de 2022, o estado atingiu um valor de R\$1.162.249, e em 2023, esse valor foi de R\$744.840. Esses dados evidenciam a significativa importância dessa cultura na geração de renda e emprego para muitas famílias, além de seu impacto positivo na economia regional e nacional.

TABELA 2 - PRODUÇÃO DE FOLHA VERDE DE ERVA-MATE, 2019-2021

UF	Área (ha)			Produção (t)			Valor (Milhões de R\$)		
	2019	2020	2021	2019	2020	2021	2019	2020	2021
Paraná	26.476	29.876	30.368	197.352	228.382	238.110	352.23	434.12	448.8
Santa Catarina	15.636	14.687	13.876	89.909	83.233	76.763	112.04	91.57	94.36
Rio Grande do Sul	28.022	27.413	27.699	233.434	214.552	242.018	327.6	355.19	356.35
Mato Grosso do Sul	171	146	163	1.564	1.379	1.096	0.84	3.17	1.33
Total	70.305	72.122	72.106	522.259	527.546	557.987	792.71	884.05	900.84

Fonte: SIDRA/IBGE (2022).

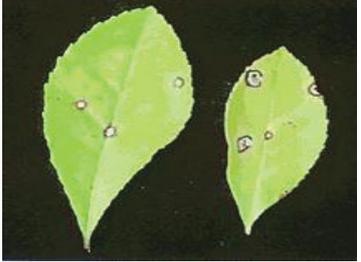
Embora a cultura da erva-mate tenha uma grande importância financeira na economia local, as plantações estão enfrentando desafios crescentes devido à diminuição da produção e à morte das plantas, causadas por várias doenças, como a cercosporiose, pinta-preta, fumagina, fuligem, podridão-do-tronco, nematoide, podridão das raízes (Piassetta et al., 2021). Muitas dessas doenças atacam simultaneamente as folhas da cultura e se propagam rapidamente, o que é agravado pelas condições da região, apesar de existir esforço por parte dos produtores para reduzir os impactos dessas doenças por meio da implementação de métodos tradicionais, como o uso de pesticidas e inseticidas. O uso excessivo desses produtos pode afetar negativamente o meio ambiente e a saúde humana (Ali et al., 2022). Além disso, o diagnóstico incorreto das doenças pode levar a um uso inadequado de pesticidas e inseticidas, o que não apenas aumenta os custos para os produtores, mas também agrava os problemas de resistência de pragas e ambiental.

O Quadro 1 apresenta as principais doenças que afetam a cultura de erva-mate, destacando características específicas como tombamento (damping off), mancha-da-folha (pinta-preta) e antracnose. Além destas, são mencionadas doenças

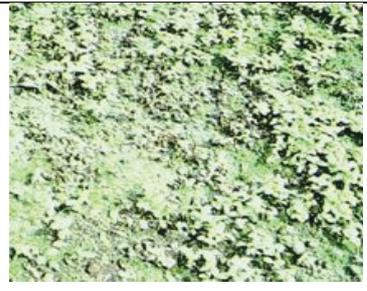
de importância secundária, tais como podridão-das-raízes, cercosporiose, nematóides, e outras como fumagina, fuligem, podridão-do-tronco e queda-de-folhas (Poletto et al., 2010; Marques et al., 2013; Grigoletti Júnior e Auer, 2013).

O controle eficaz e a gestão adequada dessas enfermidades são cruciais para assegurar uma produção saudável e sustentável de erva-mate. As imagens que ilustram diversas doenças, incluindo Pinta-preta, Antracnose, Cercosporiose, Fumagina, Fuligem, Podridão-do-tronco, Podridão das raízes, Nematóide e Queda-de-folhas, conforme detalhado no Quadro 1, foram obtidas do estudo realizado por Grigoletti Júnior e Auer (2013). Já as imagens que evidenciam os sintomas da Requeima foram registradas pelo próprio autor durante a pesquisa na área de estudo.

QUADRO 1 - CARACTERIZAÇÃO DE DOENÇAS DE CULTURA DE ERVA-MATE, SEGUNDO GRIGOLETTI JÚNIOR E AUER (2013)

Mancha de doença	Caracterização
	<p>Pinta-preta: Causada pelo fungo <i>Cylindrocladium spathulatum</i>. Os sintomas aparecem como lesões foliares escuras, arredondadas, às vezes concêntricas, no interior ou nas bordas do limbo. As manchas podem alcançar até 2 cm de diâmetro e ocorrem em folhas adultas, provocando sua queda prematura.</p>
	<p>Antracnose: causada por <i>Colletotrichum sp.</i> Os sintomas são manchas escuras, irregulares, incidindo principalmente nas bordas e causando deformações nas folhas jovens é uma doença que ocorre principalmente nas brotações jovens, ápices, folhas e ramos jovens, impedindo seu crescimento. As condições favoráveis são o sombreamento excessivo e a umidade excessiva.</p>
	<p>Cercosporiose: Causada por <i>Cercospora sp.</i>, é considerada uma doença secundária por sua baixa incidência, sua disseminação bastante lenta e por causar poucos danos à cultura. Os sintomas são o surgimento de manchas arredondadas, pequenas, bem delimitadas, acinzentadas, com halo escuro, apresentando pequenas pontuações. As manchas são frequentes em folhas adultas.</p>

	<p>Fumagina: causada por <i>Meliola sp.</i> e <i>Capnodium sp.</i>, que colonizam superficialmente folhas e ramos. Caracteriza-se pela presença de uma crosta espessa e negra cobrindo total ou parcialmente a parte dorsal das folhas e ramos da erva-eira. Geralmente, formigas, cochonilhas e pulgões estão associados a esses sintomas. O crescimento denso e escuro sobre os órgãos clorofilados pode reduzir a atividade fotossintética das plantas atacadas, podendo levá-las à morte. É uma doença que não apresenta grande importância econômica.</p>
	<p>Fuligem: causada por <i>Asterina mate.</i> Os sintomas são observados na face ventral das folhas, onde ocorrem manchas superficiais escuras de forma circular que podem se justapor, cobrindo totalmente a folha. É uma doença de pouca importância econômica, pelos danos causados.</p>
	<p>Requeima: causada pela presença de <i>Phytophthora infestans</i> - é uma das principais ameaças fúngicas à cultura de erva-mate. Encontra-se disseminada em todas as regiões, exceto onde as condições climáticas desfavoráveis de temperatura e umidade relativa. favoráveis. Todas as plantas, independentemente da idade, são suscetíveis a esse fungo, capaz de afetar toda a parte aérea da planta. Os sintomas nas folhas podem variar, influenciados por fatores como temperatura, umidade, intensidade luminosa e resistência do hospedeiro. Inicialmente, as folhas apresentam tamanho reduzido e formato irregular, com coloração que varia de verde-claro a escuro. Um halo encharcado é visível, delimitando a lesão do tecido sadio. Em condições de alta umidade, observa-se um crescimento esbranquiçado no interior da folha, correspondente às estruturas reprodutivas do fungo. Em contraste, sob baixa umidade, o crescimento das lesões é interrompido, e o tecido torna-se quebradiço. Essa diversidade de sintomas destaca a influência direta das condições ambientais na manifestação e progressão da requeima (Kimati et al., 2005).</p>
	<p>Podridão-do-tronco: causada por fungos basidiomicetos. Os ramos podados não emitem brotações, secam, apodrecem e, muitas vezes, matam a planta. Os sintomas são causados pela dificuldade de cicatrização dos ferimentos causados pela poda, que facilita a penetração de fungos oportunistas. Geralmente, na base do tronco, aparecem frutificações típicas de basidiomicetos (cogumelos).</p>

	<p>Podridão das raízes: fungos associados são <i>Fusarium sp.</i>, <i>Rhizoctonia sp.</i> e <i>Phyium sp.</i>, pode ocorrer nas sementeiras como nas mudas repicadas. Os sintomas manifestam-se na parte aérea, na forma de manchas foliares, amarelecimento, queda de folhas, redução do crescimento, murcha e secamento das mudas. Os sintomas podem ser confundidos com os provocados por repicagem inadequada das mudas ou pela falta de água.</p>
	<p>Nematoide: causada por fungos de gênero <i>Meloidogyne</i>. Os sintomas iniciais são a paralisação do crescimento das mudas e o amarelecimento das folhas, ainda murcha e secamento a partir do ápice da muda. Nas raízes, o aparecimento de galhas revela a presença de nematóides. Pode ocorrer em viveiros, onde não é utilizada a prática de desinfestação do substrato e pelo estresse da muda ocasionado por excesso de tempo no viveiro.</p>
	<p>Queda-de-folhas: observa-se a presença de <i>Cylindrocladium spathulatum</i>, o mesmo que provoca a queda das folhas no viveiro. Os sintomas são variados, podendo ocorrer o amarelecimento e posterior queda das folhas, das totalmente verdes e a das verdes manchadas. As condições favoráveis ao surgimento dessa doença são a compactação do solo, as estiagens prolongadas e longos períodos de precipitação pluviométrica.</p>

Fonte: Grigoletti Júnior e Auer (2013).

Entretanto, no caso das doenças que afetam a cultura da erva-mate, a adoção de medidas de controle enfrenta consideráveis desafios. Isso ocorre principalmente devido à escassez de informações abrangentes relacionadas aos aspectos epidemiológicos dessas doenças, bem como à dificuldade de realizar diagnósticos precoces e precisos.

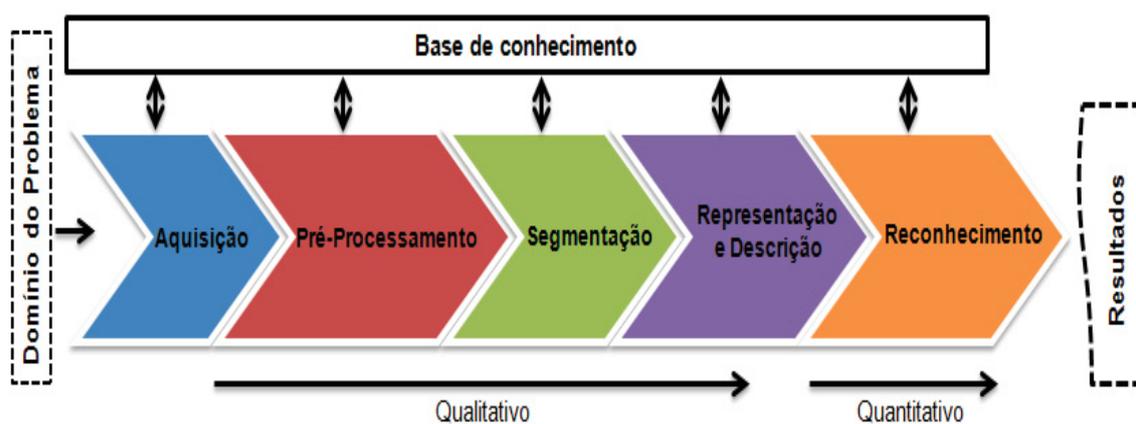
2.3 PROCESSAMENTO DIGITAL DE IMAGENS

O processamento de imagem abrange uma série de etapas de manipulação de dados em que a entrada e saída apresentam-se na forma de imagens, como fotografias ou quadros de vídeo (Young et al., 1998). A aplicação destas técnicas

permite a extração e identificação de informações contidas nas imagens, ao mesmo tempo em que se aprimora a qualidade visual de aspectos estruturais específicos. A maioria dos princípios de processamento de sinais, que são aplicados a sinais unidimensionais, pode ser estendida para o processamento de imagens (que é bidimensional), incluindo técnicas como limiarização, filtragem, redução de ruído e detecção de bordas. Por outro lado, também existem métodos desenvolvidos especificamente para o processamento de imagens, como a análise de textura e conectividade (Gonzales e Woods, 1992; Young et al. 1998).

O sistema de processamento digital de imagens é composto por um conjunto de etapas (conforme ilustrado na Figura 1) capazes de produzir um resultado com base no domínio do problema. O domínio do problema refere-se à entrada a ser analisada e o resultado é a saída dos dados interpretados. Todas as etapas intermediárias têm como objetivo resolver um problema específico.

FIGURA 1 - ETAPAS DE SISTEMA DE PROCESSAMENTO DE IMAGENS



Fonte: Adaptado do Pedrini e Schwartz (2008).

2.3.1 Aquisição

As etapas comumente empregadas incluem a aquisição de imagens, que envolve o uso de câmeras ou outros sensores capazes de gerar grades semelhantes

às imagens, sendo estas tratadas posteriormente como imagens. Um exemplo seria a geração de grades altimétricas a partir de uma varredura realizada com LiDAR.

2.3.2 Pré-Processamento

As fases iniciais de Pré-Processamento, conforme destacado por Pedrini e Schwartz (2008) e Gonzales e Woods (2010), têm como objetivo aprimorar a qualidade da imagem através da aplicação de técnicas destinadas à redução de ruído, correção de contraste e brilho, bem como à suavização de propriedades específicas da imagem.

2.3.3 Segmentação

A segmentação envolve a extração e identificação de áreas de interesse presentes na imagem. A imagem digital é subdividida em várias unidades significativas com base na similaridade e homogeneidade das características específicas relacionadas à resolução do problema (Pedrini e Schwartz, 2008).

Várias abordagens de segmentação são apresentadas na literatura, como a: Técnica baseada em região, onde os pixels relacionados a um objeto são agrupados (por exemplo a técnica K-means ou Watershed e Fuzzy); Técnica de detecção de borda (exemplo técnica de Canny, Sobel, Gradient, Laplacian, Robert, etc); Técnica de limiarização usando informações de cores, a intensidade da cor ou o brilho do espaço de cores HSV (Iqbal and Talukder, 2020; Sobiya et al., 2022); Segmentação baseada em textura, que envolve a identificação de padrões de textura na imagem, como a matriz de coerência de níveis de cinza ou filtros Gabor (Parida e Bhoi, 2017); Segmentação espectro-espacial como o método *Simple Linear Iterative Clustering* (SLIC) baseada em superpixels agrupando diferentes pixels em imagens com base

na semelhança de cores e sua posição em um plano de imagem (Phan et al., 2022); Segmentação baseada em alongamento de contraste - amplia-se a faixa de intensidade dos pixels para destacar objectos ou regiões de interesse em relação ao fundo e uma limiarização pode ser aplicada para separa objectos do fundo com base nas intensidades modificadas (Barbedo, 2016); Segmentação com a combinação de cor abrangente com região em crescimento - começa selecionando pixels iniciais (sementes) com base em suas cores e em seguida, os pixels vizinhos com cores semelhantes são adicionados à região, formando uma área coesa de interesse (Ma et al., 2018) e; Técnicas de segmentação baseados em modelos, a considerar os seguintes exemplos: Segmentação baseada em Markov Random Field - este modelo tem restrição de suavidade da região espacial e existe correlações entre os componentes de cor (Zhang et al., 2015); Modelo de segmentação usando Gaussian Markov Random Field, onde a dependência entre os pixels são consideradas para o processo (Almakady et al., 2020); Modelo Narrow Band baseado em multirresolução (Zheng et al., 2017), entre outros.

No entanto, a desvantagem de algumas das técnicas de segmentação aqui mencionadas é que elas não são muito adequadas para cenários de imagens com fundo complexo.

2.3.4 Representação e descrição

A etapa de representação e descrição tem como objetivo extrair as características ou atributos que serão empregados na fase de classificação das regiões, visando à identificação de objetos ou de suas partes nas imagens. A tarefa principal consiste em destacar as diferenças e semelhanças entre os segmentos obtidos na etapa de segmentação, permitindo assim a identificação, rotulagem ou estabelecimento de padrões de classificação, como apontado por Pratt (2001). Nesse processo, os elementos mais comumente utilizados para a extração de características incluem cor, forma, tamanho e textura.

No que diz respeito à textura, esta é um elemento fundamental na percepção visual humana e é amplamente empregada em muitos sistemas de visão computacional. Rosenfeld e Kak (1982) definiram textura como um agrupamento de elementos similares em uma imagem, enquanto Pratt (2001) a descreveu como os padrões semi-repetitivos de pixels encontrados em cenas naturais. Portanto, a análise de textura é aplicada em uma variedade de campos e aplicações, abrangendo desde a classificação de texturas até a segmentação de imagens, a síntese de imagens e o reconhecimento de padrões.

Ao longo das últimas décadas foram propostos diversos métodos de extração de características. Cada um desses métodos apresenta suas próprias vantagens e limitações específicas. Alguns deles mantêm seu desempenho inalterado diante de operações como translação, rotação, afinidade e transformação de perspectiva; outros se destacam por sua baixa complexidade computacional; alguns são notavelmente simples de implementar, entre outras características distintas.

Humeau-Heurtier (2019) conduziu uma pesquisa abrangente sobre os métodos de extração de características de textura, e esses métodos foram categorizados em sete classes distintas, conforme descrito na Tabela 3: abordagens estatísticas, abordagens estruturais, abordagens baseadas em transformações, abordagens baseadas em modelos, abordagens baseadas em gráficos, abordagens baseadas em aprendizado e abordagens baseadas em entropia.

TABELA 3 - MÉTODOS DE EXTRAÇÃO DE RECURSOS DE TEXTURA

Classes	Métodos
Abordagens Estatísticas	<ul style="list-style-type: none"> - Matriz de co-ocorrência de nível de cinza; - Matriz de execução de nível de cinza; - Abordagens baseadas em autocorrelação; - Abordagens baseadas em padrões locais mapeados; - Histograma de magnitudes de gradientes; - Padrão de energia local; - Variogram; - Recursos de Tamura; - Padrão binário local e variantes; - Histogramas de índice de forma; - Caminhada determinística; - Descritor local de Weber;

Abordagens estruturais	- Textura considerada como a replicação - de forma mais ou menos regular
Abordagens baseadas em transformação	<ul style="list-style-type: none"> ▪ Bancos de filtro: recursos de textura da lei, uma abordagem de energia de textura ▪ Abordagens baseadas em quatro transformas ▪ Abordagens baseadas na decomposição de Gabor ▪ Abordagens baseadas em wavelet ▪ Abordagens baseadas em tesouras ▪ Histograma de recurso de transformar localmente codificado (letrist) ▪ Abordagens baseadas em contorno
Abordagens baseadas em modelo	<ul style="list-style-type: none"> ▪ Abordagem complexa baseada em rede ▪ Modelos de mosaico ▪ Modelos de estrutura celular ▪ Modelos de cobertura ▪ Modelos de campo aleatórios ▪ Modelos auto-regressivos ▪ Modelos médios móveis ▪ Modelos de textura de campo Markov aleatório ▪ Modelo generalizado de correlação longa ▪ Medidas de textura baseadas no fractal ▪ Modelos gravitacionais ▪ Wold Decomposition
Abordagens baseadas em gráficos	<ul style="list-style-type: none"> ▪ Estruturas de gráficos locais ▪ Gráfico da abordagem de caminhada turística ▪ Menores caminhos em gráficos
Abordagens baseadas na aprendizagem	<ul style="list-style-type: none"> ▪ Métodos de aprendizagem de vocabulário ▪ Métodos baseados em máquina de aprendizagem extrema ▪ Métodos de aprendizagem profunda
Abordagens baseadas na entropia	<ul style="list-style-type: none"> ▪ Entropia de amostra dimensional ▪ Entropia de distribuição bidimensional ▪ Entropia multiescalas duas dimensionais

Fonte: Humeau-Heurtier (2019).

De maneira semelhante, podem ser utilizados outros descritores baseados na textura para contribuir nas tarefas de identificação e classificação de doenças em plantas. Exemplos do uso de descritores de textura são encontrados na literatura, como aqueles usando a Matriz de Coocorrência de Nível de Cinza - GLCM (Fulari, et al., 2020; Saputra et al., 2020; Abdu et al., 2020; Yogeshwari e Thailambal, 2021; Ahmed e Yadav, 2023), o Filtro de Gabor (Kaur, 2021; Vishnoi et al., 2022) e o Padrão Binário Local - LBP (Medhi et al., 2016; Aruraj et al., 2019; Barburiceanu et al., 2020).

Autores como Andrearczyk (2017), Andrearczyk e Whelan (2016), e Andrearczyk e Whelan (2018) propuseram a exploração do uso de Redes Neurais Convolucionais (CNNs) na análise de textura. Lin e Maji (2016) conduziram um estudo sobre representações de textura baseadas em CNN, enquanto Cimpoi et al. (2016) demonstraram que o aprendizado profundo pode ser aplicado em diversas áreas da análise de textura. Esses autores relataram que os modelos investigados se mostraram eficazes na síntese de textura e na manipulação de conteúdo de imagens usando atributos de textura. Suas pesquisas contribuíram para o avanço da compreensão e aplicação de CNNs na análise e processamento de textura em imagens.

A extração de atributos ou propriedades é de extrema importância para um classificador automático. É essencial escolher as características de forma adequada, de modo que elas garantam alto grau de separabilidade entre as diferentes classes. Os descritores podem ser categorizados em dois grupos: globais e locais. Os descritores globais são calculados a partir de características convencionais, como cor, textura e forma. A extração de características baseada nas cores é amplamente usada para identificar e extrair objetos em imagens, sendo reconhecida como uma das características mais significativas para o reconhecimento visual (Chaki e Dey, 2020; Abdullah et al., 2021).

Os descritores locais são computados em regiões próximas a pontos de interesse na imagem e envolvem algoritmos altamente especializados. Alguns exemplos desses descritores locais incluem *Scale Invariant Feature Transform* (SIFT), *Speeded Up Robust Features* (SURF), HAAR (*Haar Wavelet*), Histograma de um Gradiente Orientado (HOG), *Oriented FAST and Rotated BRIEF* (ORB), entre outros.

Entende-se por métodos artesanais de extração de características, no contexto deste trabalho, a proposta de descritores pelo usuário ou obtidos na literatura, desenvolvidos por seres humanos com base no seu conhecimento a respeito do problema de interpretação de imagens. Mais recentemente, características podem ser calculadas por um sistema computacional automático. Estas características podem ou não ter significado para o ser humano, mas servem para a

solução do problema. Este tipo de extração de características caracteriza o aprendizado de máquina e é baseado na análise de um grande conjunto de amostras significativas.

2.3.5 Classificação

O reconhecimento ou classificação representa a última etapa do processamento de imagens, na qual atribui-se um identificador ou rótulo aos objetos ou regiões com base nas características fornecidas por seus descritores. Esse processo de interpretação tem como objetivo atribuir significado ao conjunto de objetos reconhecidos (Pedrini e Schwartz, 2008). Além disso, de acordo com Pedrini e Schwartz (2008, p.397), a classificação de padrões visa estabelecer um mapeamento que relaciona as propriedades extraídas das amostras a um conjunto de rótulos, com a restrição de que amostras com características semelhantes sejam mapeadas para o mesmo rótulo. Em essência, a classificação busca agrupar objetos ou regiões em categorias ou classes com base em suas características distintivas, tornando possível a identificação e rotulação automática dos elementos presentes na imagem.

De acordo com Ramaniharan et al. (2016), Kumar et al. (2018), Kumar et al. (2020), Donkor et al. (2022), os classificadores podem ser estatísticos paramétricos e não paramétricos. Os classificadores estatísticos paramétricos, amplamente usados no processamento de imagens e no reconhecimento de padrões, assumem uma distribuição específica para os dados e a suposição de independência entre as características, dada a classe, simplificando o modelo. Exemplos de classificadores paramétricos incluem o Teorema da Probabilidade de Bayes, que se baseia na probabilidade de um evento ocorrer com base em outro evento conhecido a priori, e a Regressão Logística. São eficientes e adequados para dados bem comportados. Por outro lado, os classificadores não paramétricos, como redes neurais, não fazem suposições específicas sobre a distribuição dos dados. Eles aprendem

automaticamente com os dados, sem a necessidade de especificar regras ou características explicitamente. Essa flexibilidade torna os classificadores neurais adequados para lidar com problemas complexos de classificação nos quais os padrões podem ser difíceis de definir explicitamente. Exemplos de algoritmos não paramétricos incluem *Support Vector Machine*, *K-nearest Neighbor*, *Random Forest* e redes neurais profundas.

A classificação pode ser categorizada com base no método de treinamento dos algoritmos em duas principais abordagens: supervisionada e não supervisionada. Além disso, pode-se destacar também aprendizagem por reforço. Na classificação supervisionada, são utilizadas imagens previamente classificadas com o objetivo de orientar o classificador, indicando antecipadamente as características a serem observadas. Alguns dos algoritmos de classificação supervisionada mais populares incluem o *Support Vector Machine* (SVM), desenvolvido por Vapnik, segundo Qi et al. (2021), o *Random Forest* (RF), proposto por Ho em 1995 (Ho, 1995), e o *K-nearest Neighbor* (KNN). Cada um desses algoritmos tem suas próprias características e aplicações específicas, tornando-os valiosos recursos na tarefa de classificação de dados, incluindo imagens. Na classificação não supervisionada, o processo de aprendizado ocorre ao longo do tempo à medida que os algoritmos são aplicados a um maior número de imagens, sem a necessidade de rótulos prévios (Pedrini e Schwartz, 2008, pp. 398).

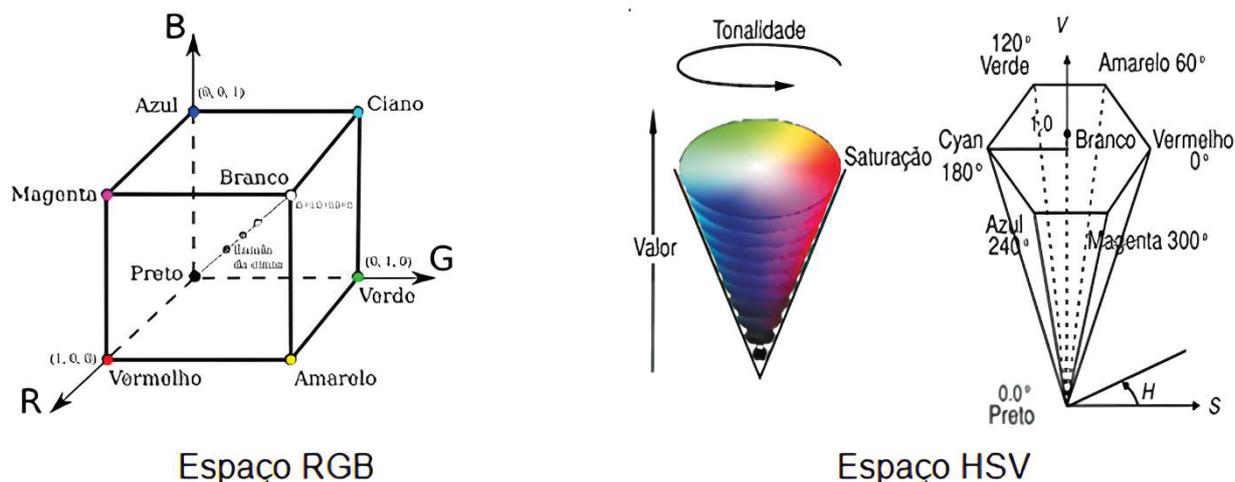
A base de conhecimento desempenha um papel fundamental no sistema de processamento de imagens. Neste contexto, a base de conhecimento assume a crucial função de descrever e armazenar as informações coletadas em cada etapa da resolução do problema (Gonzales e Woods, 2010). Atua como o elo condutor na comunicação entre as diversas etapas do processamento de imagens, sendo fundamental para a consecução de tarefas específicas. Portanto, a abrangência e complexidade da base de conhecimento podem variar consideravelmente, uma vez que são dependentes da aplicação em questão. Nesse sentido, a escolha criteriosa e a qualidade dos descritores desempenham um papel central na obtenção de resultados satisfatórios no processo de reconhecimento ou classificação. A determinação de um número adequado de atributos pode ser um fator determinante

para o sucesso na etapa de aprendizado. Por outro lado, a utilização de um espaço de descritores muito limitado pode revelar-se insuficiente para uma representação adequada dos objetos ou classes de interesse, resultando em um aumento considerável no número de erros no processo de reconhecimento.

2.4 SISTEMAS DE CORES

Nas imagens coloridas, as cores são representadas usando três bandas por imagem, geralmente as componentes no sistema de cores RGB (*Red, Green e Blue*). No sistema RGB, amplamente adotado, cada pixel é descrito por um vetor tridimensional, com a intensidade das cores primárias - vermelho (R), verde (G) e azul (B) - $f(x, y)=[R,G,B]$, como documentado por Stipanicev (1994) e Koschan e Abidi (2008). A visualização de imagens no padrão RGB, é possível pela combinação aditiva de cores (luz), seguindo o modelo natural da visão humana. Esse sistema pode ser representado como um sistema de coordenadas cartesianas, conforme ilustrado na Figura 2. As componentes do sistema RGB podem ser transformadas em outros sistemas, dentre os que se destaca o sistema HSV (Hue, Saturation, Value), devido a que suas componentes são facilmente compreendidas pelo homem (Stipanicev, 1994). A tonalidade (H - Hue) identifica o comprimento de onda dominante da cor, a saturação (S - saturation) mede a pureza da cor, indicando a proporção da cor presente, e a componente valor (V - Value) está associada ao brilho, que denota a intensidade da luz contida na cor. O modelo HSV é detalhadamente descrito nos estudos de Munsell e Ostwald (Bond & Nickerson, 1942).

FIGURA 2 - REPRESENTAÇÃO ESQUEMÁTICA DO ESPAÇO RGB E HSV



Fonte: Adaptado do Stipanicev (1994).

As imagens coloridas desempenham um papel crucial em diversas áreas, principalmente no campo do processamento de imagens, onde informações essenciais são extraídas. No entanto, é importante notar que muitos métodos recentemente desenvolvidos para o processamento de imagens, como detecção de bordas, filtragem e aprimoramento de imagens, tradicionalmente se concentram em imagens monocromáticas. Com os avanços tecnológicos, os dispositivos de captura de cores tornaram-se mais padronizados e são amplamente empregados em diversas aplicações, incluindo captura e processamento de imagens, visão computacional, vigilância, satélites e aplicações aéreas nos setores civil, militar e espacial. A representação de cores predominante é o sistema RGB (vermelho-verde-azul), que é amplamente adotado em praticamente todas essas áreas. Destaca-se a existência de vários sistemas de cores, que incluem XYZ, RGB, HSV e HLS (Burger e Burge, 2022; Skala, 2023). No contexto da detecção e classificação de doenças em plantas, foco principal da presente pesquisa, a atenção se volta para os padrões de cores convencionais, como o RGB, e para o espaço de cores HSV, conforme discutido nos estudos de (Madiwalar et al., 2017; Devaraj et al., 2019; Chaki e Dey, 2020; Abdu et al., 2020).

Para Stipanicev (1994) e Burger e Burge (2022), um dos pontos fracos do modelo RGB e, em parte, do modelo HSI para especificar imagens coloridas é sua

natureza não uniforme (Figura 2- lado direito). Distâncias iguais no espaço de cores RGB geralmente não correspondem à diferença igual na percepção das cores. Por causa disso, alguns outros modelos de cores são propostos. Um deles é o modelo de cores YIQ do padrão americano NTSC para sinal de vídeo composto, que é essencialmente uma transformação linear do modelo RGB com informações de luminância codificadas no componente Y e crominância em I e Q. De acordo com Gonzales e Woods (2010), a luminância refere-se à quantidade de cores pretas e brancas presentes na imagem e a crominância diz respeito a tonalidade da cor, a frequência dominante do raio de luz.

2.5 INTELIGÊNCIA ARTIFICIAL

A inteligência artificial (IA) é um campo científico que confere às máquinas a habilidade de aprender sem a necessidade de programação estrita. De forma mais abrangente, a IA engloba uma variedade de técnicas que permitem que os computadores imitem o comportamento humano, podendo assim tomar decisões independentes ou com intervenção mínima humana para resolver tarefas complexas (Russell e Norvig, 2021).

A IA se fundamenta em modelos analíticos que têm a capacidade de gerar previsões, regras, respostas ou resultados semelhantes (Janiesch et al., 2021). Exemplos desses modelos incluem o raciocínio baseado em casos, sistemas baseados em regras (Cihan, 2020), algoritmos genéticos (Lambora et al., 2019; Zivkovic, et al., 2021), modelos fuzzy (Avdeenko e Makarova, 2017; Cihan, 2020) e sistemas multiagentes (Cardoso & Ferrando, 2021). As primeiras tentativas de construir modelos analíticos envolviam a programação explícita de relações conhecidas, procedimentos e lógica de decisão em sistemas inteligentes, através da criação de regras artesanais (como os sistemas especialistas para diagnósticos médicos) (Goodfellow et al., 2016; Russell e Norvig, 2021). No entanto, com o avanço das estruturas de programação, a disponibilidade de dados (graças às tecnologias de

big data) e o amplo acesso ao poder computacional necessário, os modelos analíticos são cada vez mais construídos utilizando o que comumente se chama de Aprendizado de Máquina (Machine Learning- ML) (Brynjolfsson & McAfee, 2017; Goodfellow et al., 2016).

O aprendizado de máquina engloba modelos supervisionados, que são comumente aplicados a problemas de regressão, classificação e agrupamento, fazendo uso frequentemente da unidade de processamento referida por Rosenblatt (1958) como "*perceptron*". As Redes Neurais Artificiais (RNAs) podem ser categorizadas em duas principais classes: as RNAs tradicionais, que consistem em apenas uma camada oculta, e as RNAs profundas, que possuem múltiplas camadas ocultas. As RNAs profundas também são conhecidas como Aprendizado Profundo (em inglês DNNs), conforme destacado por Lecun (2015).

Dependendo do problema apresentado e dos dados disponíveis, é possível identificar três categorias principais de ML, que inclui a aprendizagem supervisionada, aprendizagem não supervisionada e aprendizagem por reforço. Dentro desse campo, existem diversas classes de algoritmos, cada uma delas com várias especificações e variantes. Essas classes incluem modelos de regressão, algoritmos baseados em instância, árvores de decisão, métodos bayesianos e RNAs. A família das RNAs é especialmente de grande interesse, uma vez que sua estrutura flexível permite adaptações para uma ampla variedade de contextos em todas as três categorias de ML (Minar e Naher, 2018; Janiesch et al., 2021).

De acordo com Minar e Naher (2018) e Janiesch et al. (2021), as RNAs são concebidas com base no princípio do processamento de informações em sistemas neurais biológicos. Elas usam representações matemáticas de unidades de processamento interconectadas, os neurônios artificiais. Similar às sinapses no cérebro, cada conexão entre neurônios transmite sinais cuja intensidade pode ser amplificada ou reduzida por meio de um peso que é constantemente ajustado durante o processo de aprendizado. Os sinais são processados apenas pelos neurônios subsequentes se ultrapassarem um determinado limite, o qual é definido por uma função de ativação. Normalmente, os neurônios são organizados em redes com

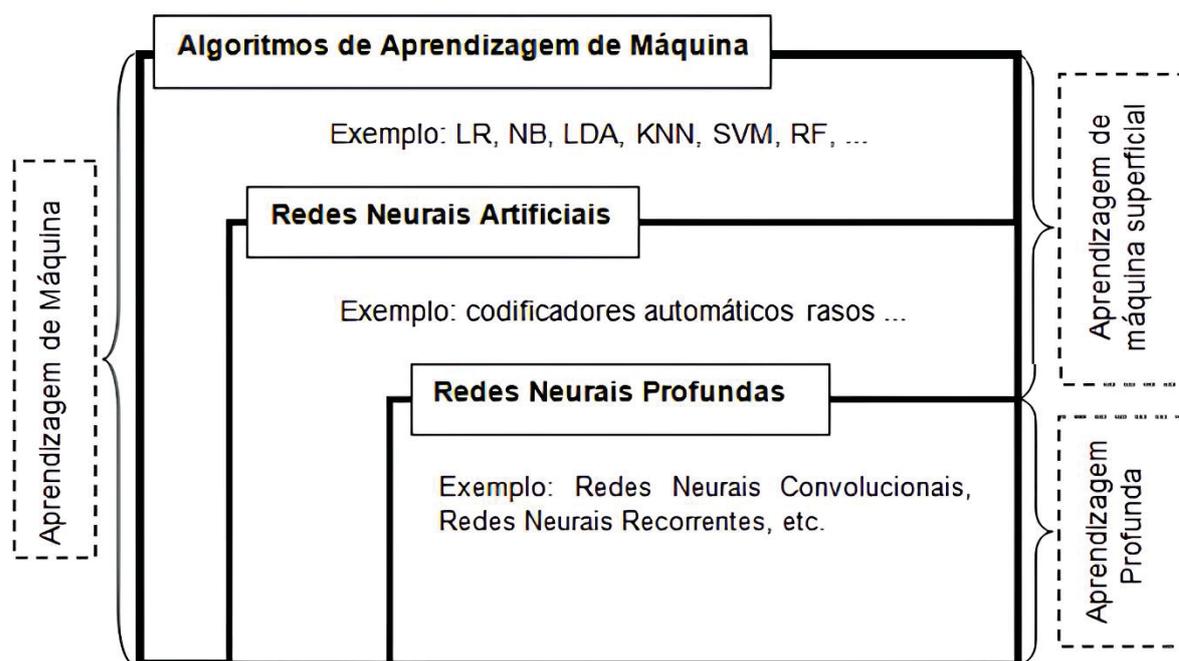
diferentes camadas. Uma camada de entrada geralmente recebe os dados de entrada, como imagens, enquanto uma camada de saída gera o resultado desejado, como detecção ou classificação. Entre essas camadas, existem as camadas ocultas, que têm a responsabilidade de aprender um mapeamento não linear entre a entrada e a saída (Goodfellow et al., 2016; Minar e Naher, 2018; Janiesch et al., 2021). O número de camadas, neurônios e outras propriedades, como a taxa de aprendizado ou a função de ativação, não podem ser aprendidos pelo algoritmo de aprendizado. Eles representam os hiperparâmetros de um modelo e devem ser definidos manualmente ou ajustados por meio de um processo de otimização (Goodfellow et al., 2016).

O conceito de Redes Neurais Profundas ou DNNs descreve a técnica de empilhar várias camadas de redes neurais para processar imagens de maneira que a saída de uma rede seja usada como entrada para a seguinte camada. Cada uma dessas redes é composta por neurônios que analisam regiões da imagem, permitindo a dedução de padrões espaciais e de cor nas imagens. Com isso, elas calculam descritores locais de características. Em outras palavras, essas redes utilizam operações avançadas, como convoluções, ou múltiplas ativações em um neurônio, em vez de empregar uma função de ativação simples. Essas características permitem que as redes neurais profundas sejam alimentadas com dados brutos de entrada e descubram as representações de características necessárias para a solução de um problema com base na fase de aprendizado. Essa capacidade central das redes neurais profundas é amplamente conhecida como aprendizado profundo. Portanto, as RNAs simples, como os autoencoders rasos, e outros algoritmos de ML, como árvores de decisão, podem ser classificados sob o termo "aprendizado de máquina superficial", uma vez que não oferecem essas funcionalidades avançadas (Minar e Naher, 2018; Janiesch et al., 2021).

Aprendizado Profundo (*Deep Learning -DL*) mostra-se particularmente eficaz em domínios com grande volume de dados e alta dimensionalidade, razão pela qual as redes neurais profundas superam algoritmos de ML rasos na maioria das aplicações que envolvem o processamento de dados de texto, imagem, vídeo, voz e áudio (LeCun et al., 2015). No entanto, em situações em que a entrada de dados

possui baixa dimensionalidade, especialmente quando há uma quantidade limitada de dados de treinamento disponíveis, o ML superficial ainda pode produzir resultados superiores (Zhang e Ling, 2018; Janiesch et al., 2021). Além disso, os resultados obtidos com algoritmos de ML superficial tendem a ser mais interpretáveis do que os gerados por Redes Neurais Profundas (Rudin, 2019; Janiesch et al., 2021). A Figura 3 representa o diagrama de classificação da IA.

FIGURA 3 - DIAGRAMA DE VENN DE ALGORITMOS DE APRENDIZADO DE MÁQUINA



Fonte: Adaptado do Goodfellow et al. (2016, p.9).

De acordo com Russakovsky et al. (2015), um dos marcos significativos que impulsionaram o rápido avanço das Redes Neurais Profundas foi o desafio ImageNet *Large Scale Visual Recognition Challenge* (ILSVRC). Esse desafio possibilitou que diversos laboratórios de pesquisa explorassem técnicas de Aprendizado Profundo, com o objetivo de aprimorar o reconhecimento de imagens no conjunto de dados ImageNet. Os participantes foram avaliados em três aspectos cruciais: detecção de objetos, localização de objetos e classificação de objetos presentes nas imagens. O conjunto de dados *ImageNet* é notável por sua vasta coleção, composta por milhões de imagens abrangendo uma ampla variedade de categorias, totalizando mil classes distintas que precisam ser reconhecidas.

2.5.1 Aprendizado de Máquina no Diagnóstico de Doenças de Plantas

Nos últimos anos, houve avanço na aplicação de técnicas de inteligência artificial na área de visão computacional, especialmente em tarefas de análise e reconhecimento de imagens. Esse avanço tornou-se essencial devido às limitações das abordagens tradicionais, que dependem da inspeção visual e enfrentavam desafios significativos, como previsões imprecisas e problemas em análises de regressão. Esses obstáculos motivaram um progresso significativo no campo da ciência, especialmente no que diz respeito ao desenvolvimento de soluções tecnológicas para o diagnóstico preciso de doenças em plantas.

Vários algoritmos convencionais de aprendizado de máquina populares para desenvolvimento automatizado de modelos para diagnósticos de doenças de plantas, são empregados, em destaque os algoritmos: RNA - Rede Neural Artificial (Shah & Jain, 2019; Gupta & Singh, 2019; Subramani et al., 2019), PCA - Análise de Principais Componentes (Prajapati et al., 2017; Dhinesh & Jagan, 2019; Kaur & Oberoi, 2020), RL - Regressão Logística (Chen, 2020), LDA - Linear Discriminant Analysis (Zapolska et al., 2020; Yogeswararao et al., 2022), NB - Naive Bayes (Xu, 2018; Mohanapriya e Balasubramani, 2019), KNN - KNeighbors (Türkoğlu e Hanbay, 2019; Rehman et al., 2019; Abdu et al., 2020; Hatuwal et al., 2020), SVM - Support Vector Machines (Huang et al., 2018; Zhang et al., 2019; Hatuwal et al., 2020; Xian and Ngadiran, 2021), RF - Random Forest (Govardhan & Veena, 2019; Hatuwal et al., 2020; Simón et al., 2022), entre outros. Esses algoritmos desempenharam um papel fundamental no desenvolvimento de modelos de diagnóstico de doenças em plantas, contribuindo para avanços significativos na área de agricultura e na detecção precoce de problemas fitossanitários.

As pesquisas iniciais no campo da agricultura exploraram a espectroscopia como uma ferramenta valiosa para identificar doenças em plantas, e avaliando o vigor da germinação de sementes por meio da análise das frequências espectrais de suas amostras. Plantas e sementes infectadas exibem características distintas em comparação com aquelas saudáveis. Isso se deve, em parte, às diferenças na

quantidade de luz absorvida na faixa do infravermelho próximo (ou Near Infrared - NIR), o que possibilita a detecção de anomalias. Um exemplo notável desse enfoque foi apresentado por Zhang et al. (2012), quando utilizaram dados espectrais para identificar a doença de oídio em plantações de trigo. Para isso, empregaram testes de correlação, regressão e teste t para amostras independentes na análise de 32 características espectrais distintas extraídas das amostras. Os resultados revelaram que ainda existem desafios significativos a serem superados para o diagnóstico eficaz de doenças em plantas (Jayakumar, 2019). Medeiros et al. (2020) conduziram uma pesquisa que se concentrou na aplicação da espectroscopia de infravermelho próximo transformada de Fourier e técnicas de imagem de raios-X para prever o vigor de germinação de sementes e os resultados foram promissores.

Atualmente, são introduzidas câmeras de alta tecnologia que permitem aprimorar a detecção e análise de problemas nas plantas. Por exemplo, as câmeras, multiespectrais, assim como hiperespectrais, possuem filtros de alta qualidade projetados especificamente para capturar uma ampla gama de ondas refletidas. Dois estudos notáveis, conduzidos por Lu et al. (2017) e Zheng et al. (2019), utilizaram dados hiperespectrais em suas pesquisas. O primeiro concentrou-se na detecção de folhas de morango infectadas por antracnose, empregando algoritmos como o KNN (K-Nearest Neighbors) e análise discriminante de Fisher (FDA). O segundo estudo investigou a cultura de trigo e observou mudanças nos dados espectrais e em índices como o Índice de Refletância Fotoquímica (IRF) e o Índice de Refletância de Antocianina (IRA) ao longo do desenvolvimento da planta. Ambos os estudos destacaram fatores que afetaram o desempenho dos métodos adotados, incluindo desafios no processo de coleta de dados, a falta de experiência na captura de imagens e a interferência de sombras causadas pela luz solar. Outros estudos, como os de Li et al. (2018) e Zhang et al. (2020), também empregaram sensores multiespectrais e hiperespectrais. O primeiro estudo teve como objetivo a detecção precoce de hematomas em pêssegos, enquanto o segundo visava prever mudanças na qualidade das sementes de trigo. Como resultado, os autores enfatizaram a importância de explorar tecnologias alternativas com o objetivo de aprimorar a precisão de suas abordagens.

Embora tenha havido avanços no uso de imagens para tarefas de identificação e classificação, como é o caso de diagnóstico de doenças em plantas, a precisão destes métodos tem sido objeto de pesquisa. Na implementação dessas técnicas, autores como Francis et al. (2016) e Dhaware e Wanjale (2017) destacaram a importância do processo de coleta e processamento de imagens, particularmente a segmentação, como elementos cruciais para a melhoria do desempenho dos modelos. De acordo com Padol e Yadav, (2016), Zhang et al. (2017), Chen et al. (2020), as taxas de reconhecimento de doenças dependem fortemente da segmentação da lesão, e as características são frequentemente projetadas artificialmente. Isso requer trabalho árduo, conhecimento especializado e, em certa medida, envolve uma subjetividade considerável, uma vez que não é fácil determinar quais características são ideais para extrair no processo de diagnóstico de doenças, especialmente dadas as dificuldades associadas à coleta de imagens em condições de luz natural não controlada, que podem apresentar características que tornam a análise da imagem desafiadora. Além disso, condições de fundo complexas frequentemente contribuem para a dificuldade na segmentação precisa da lesão na imagem (Barbedo, 2016; Liu e Wang, 2021). A segmentação de folhas em condições reais de campo apresenta desafios devido a fatores como iluminação variável, sombras (Sharma e Gupta, 2020; Hasan et al., 2022), interseções entre as folhas e outras plantas, oclusão das folhas, fundo complexo (Liu et al., 2022), e as características dos sintomas são representadas por múltiplas lesões de cores variadas na mesma folha em diferentes estágios de infecção (Hasan et al., 2022), entre outros.

2.5.2 Aprendizado Profundo e Técnicas de Transferência de Aprendizado

O pré-processamento de imagens, a segmentação e a extração de características são elementos cruciais nos sistemas de detecção automática, desempenhando um papel fundamental na obtenção de uma classificação precisa e eficiente (Janiesch et al., 2021; Liu e Wang, 2021). Portanto, a seleção apropriada dessas técnicas exerce um impacto significativo no desempenho do modelo, e essa

escolha pode depender de uma variedade de fatores, incluindo o tipo de objeto que se deseja extrair, as características da imagem, a complexidade da cena e outros aspectos relevantes.

2.5.2.1 Detecção e segmentação

A vasta lista de modelos de aprendizagem profunda para detecção de objetos inclui detectores de dois estágios, como R-CNN (Rede Neural Convolutacional Baseada em Região), Fast R-CNN, Faster R-CNN, Mask R-CNN e aqueles que envolvem um estágio, a saber: SSD (Single Shot Detector), RetinaNet, YOLO (You Only Look Once) e suas extensões como YOLOv2, YOLOv3, YOLOv4, YOLOv5, para citar alguns.

Tanto os detectores de um como de dois estágios apresentam vantagens e desvantagens (Liu et al., 2020). Detectores de dois estágios, como R-CNN (Girshick et al., 2014), Fast R-CNN (Girshick, 2015), Faster R-CNN (Girshick, 2015) e Mask R-CNN (He et al., 2017), treinam redes neurais convolucionais (CNNs) ponta a ponta para classificação de regiões (Liang et al., 2015). Primeiro, os detectores de dois estágios, geram um conjunto de locais onde o objeto é suspeito de ser encontrado (regiões candidatas) através de uma Rede de Região Proposta (RPN) no mapa de características. Estes tipos de detectores classificam cada região candidata como um “objeto” ou “plano de fundo” enquanto executam simultaneamente a regressão de localização (Girshick et al., 2014; Erhan et al., 2014; Szegedy et al., 2014; Li, 2021). Os detectores de dois estágios melhoraram de várias maneiras (precisão, velocidade, desempenho, etc.). Porém, alguns problemas persistem, como o tempo necessário para treinar a rede, o algoritmo de busca seletiva gerando propostas para regiões ruins e o fato de não ser eficiente quando implementado em tempo real.

Redes Neurais Convolucionais Baseadas em Região (RCNN) e suas variantes mais rápidas extraem muitas propostas de região da imagem de entrada e usam uma CNN para realizar a propagação direta em cada proposta de região para extrair descritores e então usar esses descritores para prever a classe e a caixa

delimitadora de cada proposta. Seu sucesso se deve em grande parte à transferência de representações de imagens pré-treinadas supervisionadas para classificação de imagens para detecção de objetos. Esses métodos, no entanto, requerem uma passagem direta pela rede convolucional para extrair características de cada objeto proposto, levando a uma carga computacional pesada (Girshick et al., 2014, Jiang e Learned-Miller, 2017).

Do R-CNN ao Faster R-CNN, muitas melhorias ocorreram. Uma das principais melhorias do R-CNN para o Fast R-CNN é que a propagação direta da CNN é realizada em toda a imagem, em vez de alimentar cada região distorcida da imagem proposta para a CNN. Também introduz a camada de agrupamento de regiões de interesse, para que os descritores possam ser extraídos da mesma maneira para regiões de interesse, mesmo que tenham formatos diferentes.

Faster R-CNN, proposto por Ren et al. (2015), é uma melhoria do Fast R-CNN. Ele usa uma nova Rede de Região Proposta (RPN) para gerar regiões propostas, o que economiza tempo em comparação com algoritmos anteriores, como pesquisa seletiva. Ele usa a camada RoIPool (*Region of Interest Pooling*) para extrair um vetor de recursos de comprimento fixo de cada região proposta e realiza classificação e regressão de caixa delimitadora. Seguindo os recentes desenvolvimentos de DL na detecção e classificação de objetos, o modelo Mask R-CNN foi proposto por He et al. (2017). Mask R-CNN adiciona um terceiro ramo que gera a máscara do objeto. Além disso, a saída da máscara é diferente da saída da classe e da caixa, exigindo um layout espacial muito mais preciso para ser extraído de um objeto.

Alternativamente, têm sido utilizados detectores de um estágio, como o SSD ou a família YOLO, propostos por Joseph Redmon e Ali Farhadi em 2015, que prevêm diretamente a saída sem passar pelo estágio de proposta de região (Redmon et al., 2016). A abordagem do detector de estágio único aplica uma única rede neural à imagem completa e, em seguida, divide a imagem em regiões, prevendo caixas delimitadoras e probabilidades para cada região, ou seja, as caixas delimitadoras são ponderadas pelas probabilidades previstas.

Os algoritmos da família YOLO olham para uma imagem apenas uma vez para prever quais objetos estão presentes e onde estão (Redmon et al., 2016; Bochkovskiy et al., 2020). YOLO trata o problema de detecção de objetos como um problema de regressão. No detector YOLO, um mapa de características com amostragem reduzida é dividido em células de grade. Para cada célula da grade, camadas totalmente conectadas são treinadas para detectar objetos centralizados nesta célula usando a imagem inteira como suporte espacial. O YOLO tem pontos fracos para pequenos objetos e grupos de objetos, que se agrupam em uma única célula. Contudo, YOLOv2 (Redmon e Farhadi, 2017) é uma melhoria em relação ao algoritmo YOLO original. A principal melhoria no YOLOv2 é o uso de caixas de âncora. As caixas de âncora são um conjunto de caixas pré-definidas com tamanhos e proporções que são usadas para combinar formas prototípicas de objetos durante o treinamento. No YOLO, as caixas de âncora são usadas para prever caixas delimitadoras, prevendo as coordenadas e a classe de cada caixa de âncora. O tamanho da saída da rede é proporcional ao número de caixas de âncora por célula da grade. Portanto, as caixas de âncora são um componente crucial da arquitetura YOLO, mas não são o mesmo que as caixas delimitadoras.

O modelo YOLOv3 utiliza regressão logística para prever a pontuação de objetividade de cada caixa delimitadora (Redmon e Farhadi, 2018; Cheng, 2020). No YOLOv3, três camadas diferentes com três passadas diferentes são usadas para prever classes e posições precisas para as caixas de ancoragem. As versões YOLOv4 (Bochkovskiy et al. 2020) têm velocidade e precisão ideais em comparação com as versões anteriores de detectores de objetos. Esta versão foi especificamente projetado para eficiência em ambientes práticos, conhecidos por suas aplicações em tempo real e exigências operacionais contínuas. Essa adaptação indica que o modelo foi otimizado para oferecer desempenho eficaz e eficiente não apenas durante experimentos ou pesquisa, mas também em contextos práticos e operacionais, onde a detecção de objetos é necessária regularmente.

Outras versões como YOLOv6 (Li et al., 2022), YOLOv7 (Wang et al., 2023) e YOLOv8 (Reis et al., 2023) surgiram, visando aumentar e otimizar a velocidade e a precisão. Isto ilustra a tendência atual da pesquisa, onde a velocidade e a precisão

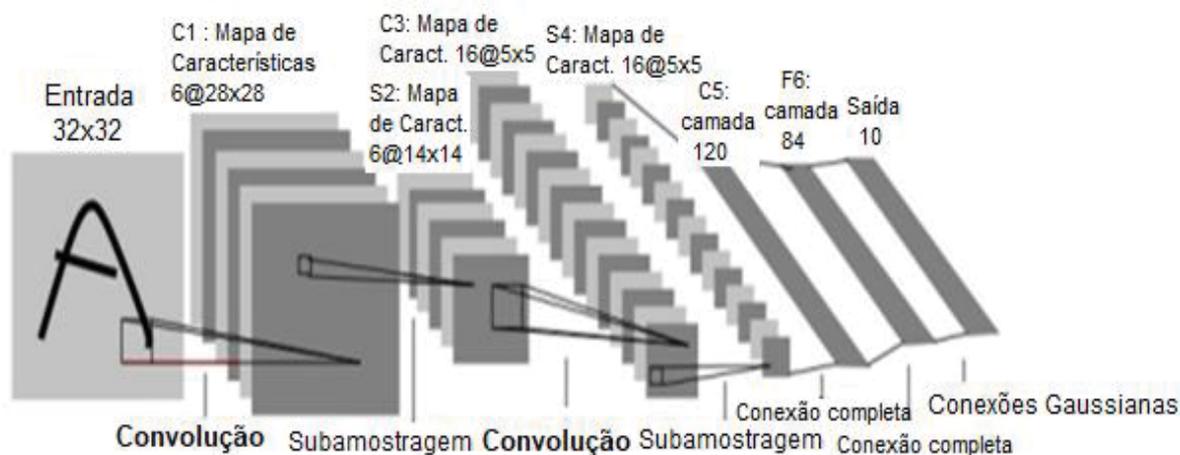
são os dois pré-requisitos para os quais os algoritmos de detecção de objetos são examinados. O YOLO é um modelo de uso geral, o que significa que pode ser utilizado para diversos fins, e a rede deve ser treinada para se adaptar a novos problemas (Redmon et al., 2016; Bochkovskiy et al., 2020).

2.5.2.2 Classificação

A baixa eficácia dos métodos tradicionais, aliada à complexidade da extração de características necessárias para o treinamento, impulsionou o desenvolvimento de técnicas avançadas para as tarefas de identificação e classificação. Nesse contexto, modelos profundos representam um conjunto inovador de algoritmos de aprendizado de máquina que demandam menos esforço nas etapas de pré-processamento. Esses algoritmos são capazes de extrair características por si mesmos, simplificando assim o processo de diagnóstico de doenças em plantas.

Neste conjunto de algoritmos, as técnicas de aprendizagem profunda, com destaque para as redes neurais convolucionais (CNNs), emergiram como a escolha preferencial para tarefas de identificação e classificação, apresentando um potencial notável para elevar as taxas de diagnóstico de doenças em plantas (Barbedo, 2018; Li et al., 2020). Além disso, elas são amplamente reconhecidas como os classificadores mais populares para o reconhecimento de imagens em problemas de grande e pequena escala (Jayakumar, 2019). As CNNs têm demonstrado uma habilidade excepcional no processamento e na categorização de imagens, como indicado por diversos estudos (Kamilaris e Prenafeta-Boldú, 2018; Kussul et al., 2017; Yalcin, 2017; Li et al., 2020).

FIGURA 4 - ESTRUTURA DE UMA CNN



Fonte: Lecun et al. (1998^a).

As CNNs apresentam uma estrutura de rede complexa que executa operações de convolução. O modelo é composto por várias camadas, incluindo a camada de entrada, a camada de convolução, a camada de pooling, a camada de conexão total e a camada de saída (conforme mostrado na Figura 4). As camadas de convolução e pooling são alternadas repetidamente, eliminando a necessidade de conexões completas entre os neurônios. Na camada de convolução é considerado um campo receptivo local, desempenha um papel crucial. Ele percorre o mapa de características para extrair informações relevantes. Posteriormente, os dados passam por várias camadas de convolução e pooling, com o objetivo de extrair características mais abstratas e relevantes. Depois desse processo, os dados entram na camada de conexão total, onde os neurônios estão completamente interconectados com os neurônios da camada superior. Por fim, os dados da camada de conexão total passam por uma etapa de classificação utilizando a função de ativação softmax, resultando na atribuição de probabilidades às classes. Esses valores são então transmitidos para a camada de saída. Esse pipeline de processamento é uma característica fundamental das redes neurais convolucionais e é responsável pelo seu desempenho notável em tarefas de reconhecimento de padrões (Lecun et al., 1998a; Liu e Wang, 2021).

As limitações das CNNs estão relacionadas principalmente à disponibilidade limitada de dados. Portanto, em problemas com conjuntos de dados pequenos, as CNNs podem enfrentar problemas de superajuste (*overfitting*), ou seja, aprender os

dados de treinamento muito bem, mas ter dificuldade em generalizar para novos dados. O uso das CNNs como extratores de descritores pode ser computacionalmente intensivo e demandar um tempo significativo de treinamento. Para lidar com esses desafios, pesquisadores como Coulibaly et al. (2019) e Jayakumar (2019) sugerem que uma abordagem eficaz envolve aumentar a quantidade de dados disponíveis e ajustar o treinamento de algumas camadas das arquiteturas com dados específicos do problema. Isso pode ajudar a mitigar as limitações inerentes ao uso das CNNs, tornando-as mais eficientes e adequadas para tarefas de processamento de imagens.

A literatura aponta outra abordagem que se destaca no processo de classificação de imagens: o uso de técnicas de transferência de aprendizado. Essas técnicas envolvem a utilização de uma rede neural previamente treinada para uma tarefa específica, abordagem diferente da pretendida como base para o treinamento na presente pesquisa. Uma rede neural pré-treinada, mesmo que tenha sido originalmente treinada para uma finalidade distinta, já possui a capacidade de compreender descritores espaciais e de cor, etapas que não precisam ser refeitas. Portanto, ao empregar uma rede neural pré-treinada, é necessário ajustar apenas uma quantidade limitada de parâmetros para resolver o problema em questão, o que resulta em um tempo de treinamento significativamente menor (Chen et al., 2022; Singh et al., 2022; Fan et al., 2022). Essas técnicas têm como objetivo principal diminuir o tempo necessário para treinar modelos de aprendizado profundo a partir do zero, ao mesmo tempo que abordam a escassez de conjuntos de dados extensos necessários para esse treinamento.

Embora a maioria das pesquisas baseadas em métodos de aprendizado profundo mencionadas no estado de arte reportarem desempenho superior, é notável que os experimentos que obtiveram níveis significativamente elevados foram conduzidos utilizando banco de dados públicos de uso geral. Esta tendência não se verifica quando as pesquisas empregam conjuntos de dados de imagens específicos ou próprios, composto para abordar problemas em contextos particulares e ambientes reais de campo de cultivo. Existem diversos fatores que podem ser identificados como responsáveis por essa discrepância. Nos bancos de dados públicos, as imagens apresentam uma diversidade limitada, frequentemente geradas em configurações

experimentais de laboratório. Isso resulta em condições menos complexas para a extração precisa e a seleção das características visíveis relacionadas à doença das plantas. Em contrapartida, as imagens capturadas em condições reais de campo de cultivo oferecem uma ampla gama de diversidade, abrangendo uma extensa variedade de características sintomáticas das doenças (Barbedo, 2018; Chen et al., 2020). Por outro lado, a complexidade do fundo nas imagens também exerce influência sobre o desempenho dos modelos. O estudo conduzido por Barbedo (2018) destacou as disparidades de precisão e acurácia ao treinar algoritmos de aprendizado profundo utilizando imagens originais sem a remoção do fundo, em comparação com imagens em que o fundo foi removido. Isso ressalta a relevância da segmentação para a tarefa de diagnóstico de patologias em plantas. Portanto, em situações em que as imagens são coletadas em ambientes não controlados, a complexidade aumenta ainda mais, uma vez que a presença de interferências e variações imprevisíveis pode comprometer significativamente o desempenho dos algoritmos automáticos de detecção.

No âmbito deste estudo, a hipótese formulada concentra-se na abordagem de questões relacionadas com as dificuldades enfrentadas ao lidar com espécies específicas fotografadas em ambientes reais de campo de cultivo, destacando os desafios encontrados no estado de arte entre os dois tipos de banco de dados (específico e público). Entre os bancos de dados públicos, merecem destaque os dados da *PlantVillage* (PVD), *Comprehensive Automation for Specialty Crops - Internal Feeding Worm Database* (CASC-IFW) e *University of California Irvine Machine Learning Repository* (UCI). Existem casos particulares de estudos, como o de Sharif et al. (2018), que combinaram um banco de dados específico com dados públicos da PVD. Isso sublinha a necessidade de uma análise detalhada do impacto dessa integração.

A utilização de um conjunto de dados primário específico oferece vantagens significativas. Ao coletar e utilizar um conjunto de dados próprio, os pesquisadores podem adaptar as soluções de detecção de patologias às condições climáticas e ambientais locais, levando em consideração os fatores que podem afetar o desenvolvimento das plantas, inclui informações detalhadas sobre as condições

ambientais, como temperatura, umidade e precipitação, que desempenham um papel crucial na saúde das plantas.

A Tabela 4 oferece uma visão geral dos diversos estudos realizados no setor agrícola, abrangendo desde aqueles que empregaram conjuntos de dados públicos e específicos, além das variadas técnicas aplicadas na segmentação de imagens de folhas com sintomas, extração de características, modelos de classificação e os respectivos resultados alcançados em cada caso. Portanto, as mesmas técnicas foram empregados em outras áreas, no caso da saúde, para modelos de diagnóstico de lesões da pele e câncer de mama.

TABELA 4 - DESCRIÇÃO DE DIFERENTES ESTUDOS COMPARATIVOS AO MODELO PROPOSTO

	Autor	Banco de dados	Espécies	Imagem de entrada, Sensor e resolução	Técnica de segmentação	Algoritmo de classificação	Extração de características	Acurácia (%)
1	Kawasaki et al. (2015)	Próprio	Pepino	800 -Imagens RGB; Tamanhos variados; Redimensionadas (224 x 224px)	Não especificado	CNN adaptada	CNN	~95,0
2	Fujita et al. (2016)	Próprio	Pepino	7.520 - Imagens RGB; Redimensionadas (224 x 224px)	Não especificado	CNN1 e CNN2	CNN	83,2
3	Pujari et al. (2016)	Próprio	Milho, Pepino, Algodoeiro e Tomate	9.912 imagens RGB; Tamanho de imagem original (1920 x1080px)	Características baseada na cor: RGB e HSI e baseado em textura: GLCM	ANN e SVM	Cor e GLCM	ANN - 87,0; SVM - 92,0
4	Suresha et al. (2017)	Próprio	Arroz	330 Imagens RGB; Tamanhos variados; Redimensionadas (255 x 255px)	Representação de espaço de cores HSV e usado método Otsu para segmentação	KNN	Característica geométrica: Área, Eixo Maior, Eixo Menor e Perímetro da parte doente das folhas	KNN- 76,59
5	DeChant et al. (2017)	Próprio	Milho	1.834- Imagens RGB; Tamanhos variados; Redimensionadas (224 x 224px)	Segmentação manual na área infectada	Pipeline de CNN	CNN	96,7
6	Ramcharan et al. (2017)	Próprio	Mandioca	11.670- Imagens RGB; Tamanhos variados	Não especificado	InceptionV3	InceptionV3	93
7	Amara et el. (2017)	Próprio	Banana	3.700 - Imagens RGB; Redimensionadas (60 x 60px)	Imagens RGB convertidas em espaço de cores HIS e LAB; Cálculo de Índice de cor de lesão	LeNet	LeNet	Imagens RGB - 92,88; Imagens cinza - 85,94
9	Zhang et al. (2017)	Próprio	Pepino e Maça	Número não especificado; Tamanhos variados; Redimensionadas (255 x 255px)	K-means, o agrupamento superpixel - primeiro a imagem foi convertida em espaço LAB e dividida em N superpixels por SLIC	C-SVM) com função de kernel de base radial	PHOG (Pirâmide de Histograma de Gradiente de Orientação)	Maça -90,43; Pepino - 92,15
10	Picon et al. (2019)	Próprio	Trigo	8.178- Imagens RGB; Redimensionadas (224 x 224px)	SLIC- clusterização em espaço de cor CIELAB	ResNet50	ResNet50	87
11	Vamsidhar et al. (2019)	Próprio	4 classes de folhas	Não especificado	RGB, espaço HIS e Clusterização com K-means	C4.5, NB, MLP, SVM	GLCM	C4.5- 89,0; NB- 86,0; MLP- 88,9; SVM-90,0; SVM (Linear Kernel) -95,6; SVM (RBF Kernel) - 94,2; SVM (Polinomial Kernel) - 95,8

12	Dhingra et al. (2019)	Próprio	Folhas de Manjerição	400- Imagens RGB; Redimensionamento não especificado	Conversão da imagem para espaço de cor CMYK e LAB	DT, NB, KNN, RF, AdaBoost, Análise Discriminante (AD), GLM	Informação de Histograma e LBP	DT -91,8; NB- 92,31; KNN- 96,92; SVM- 90,8; RF- 98,4; AdaBoost- 95,03; ANN- 80,33, AD - 80,0; GLM- 95,03
13	Nair et al. (2019)	Próprio	4 classes de folhas de milho	100- Imagens RGB; Tamanho variado; Redimensionadas (256x256px)	RGB, HSV, K-means e Usado o filtro de Gabor para ruídos	SVM e ANN	Extração de características baseadas em cores	SVM- 80,0; SVM Otimizado com GA- 86,0; ANN- 80,0; ANN Otimizado com GA- 86,0
14	Selvaraj et al. (2019)	Próprio	Banana	700 -Imagens RGB	Rotulagem manual da área infetada	ResNet, InceptionV2, MobileNetV1	ResNet, InceptionV2, MobileNetV1	Todos modelos acima de 90,0
15	Hu et al., 2019	Próprio	Chá	120- Imagens RGB	SVM	C-DCGAN +VGG16	Características de cor e de textura	90,0
16	Ma et al. (2018)	Próprio	Pepino	14.208 - Imagens RGB, Tamanho: 2592 x 1944px; Redimensionadas (800 x 600px)	Índice de excesso de vermelho (EXR), espaço de cor HSV, Padrão Binário Local (LPB) e Operações morfológicas	DCNN	CNN	93,4
17	Ramesh et al. (2018)	Próprio	Não especificado	Não especificado	Conversão RGB para HSV	RF, RL, SVM, KNN, CART, NB	HOG; Momentos Hu; Textura Haralick e Histograma de cor	RF- 70,14; RL- 65,3; SVM- 40,3; KNN- 66,76; CART - 64,7; NB- 57,61
18	Zhang et al. (2019)	Próprio	Pepino	800- Imagens RGB, Tamanho: 2456 x 2058px; Redimensionadas (240 x 240px)	Não especificado	CNN + AlexNet = GPDCNN	GPDCNN	95,18
19	Bagri and Johari (2015)	Broadtz (Público)	Categorias como tijolos, pedras, grama etc - em tom de cinza	Não especificado	Não especificado	GLCM + Momento Hu e Tamura + Momento Hu	GLCM, Momentos Hu e Textura	Tamura + Momento Hu - Recall-75,0% e GLCM + Momento Hu - Recall - 87,0
20	Lu et al. (2017)	Próprio + WWD2017	Trigo	9 230- Imagens RGB; Tamanhos variados; Redimensionadas (224 x 224px)	Não especificado	VGG-FCN-VD1 e VGG-FCN-S	VGG-FCN-VD1 e VGG-FCN-S	VGG-FCN-VD1 -97,95; e VGG-FCN-S 95,12
21	Saberi Anari (2022)	PVD + UCI	Maçã, milho, algodão, uva, pimenta e arroz	90.000- Imagens RGB; Tamanhos variados; Redimensionadas (416 x 416px)	Espaço de cor HSV	Transferência de aprendizagem	CNN	99,1
22	Sharif et al., 2018	Próprio + PVD	Citrino	580 - Imagens RGB; Tamanhos variados; 1000 - Imagens RGB - Tamanho (100x150px)	Segmentação ponderada, mapa de saliências e operações morfológicas	Método híbrido, pontuação de PCA, entropia e vetor de covariância dado M-SVM	Pontuação PCA, entropia, e vetor covariância baseado em assimetria e M-SVM	97,0 -PVD, 89-PVD + Próprio, 90,4- Próprio
23	Sharif et al. (2018)	Próprio + PVD	Cítrinos	580- Imagens RGB; Tamanhos variados; RGB- 1.000 Imagens RGB, Redimensionadas (100x150px)	Baseado no Qui-quadrado e Transformação Dimensional de Cores	Método híbrido, PCA, entropia e vetor de covariância dado M-SVM	Fusão (cor, textura e geométricas)	PVD- 97,0; PVD + Próprio- 89,0; Próprio - 90,4

24	Affi et al. (2020)	Próprio + PVD	Folha de café	54.305- Imagens RGB; Tamanhos variados	Não especificado	ResNet18, ResNet34 e ResNet50 =DAML	DAML	99
25	Chen et al. (2022)	PVD + Próprio	Amendoim, batata, maçã e 15 outras culturas	6.033- Imagens RGB; Tamanhos variados; Redimensionadas (224x224px); 54.634- Imagens RGB; Tamanhos variados; Redimensionadas (224x224px)	Não especificado	CACPNET	DNN	PVD-99,70 e Próprio - 97,7
26	Sladojevic et al. (2016)	Websites	13 Espécies de culturas	30.880 - Imagens RGB- Redimensionadas (256 x 256px)	Retângulo envolvente a área infetada	CaffeNet	CaffeNet	96,3
27	Mohanty et al. (2016)	PVD	14 Espécies de culturas	54.306 - Imagens RGB; Tamanhos variados; Redimensionadas (256 x 256px)	Máscaras geradas pela análise da cor LAB e HSB	AlexNet e GoogleNet	AlexNet e GoogleNet	99,34
28	Brahimi et al. (2017)	PVD	Tomate	14.828 - Imagens RGB; tamanhos variados; Redimensionadas (256 x 256px)	Não especificado	AlexNet, GoogleNet, SVM, RF	AlexNet e GoogleNet	GoogleNet - 99,18; AlexNet - 98,66; SVM - 94,5; RF-95,5
29	Ferentinos (2018)	PVD	25 Espécies de culturas	87.848- Imagens RGB; Tamanhos variados; Redimensionadas (256 x 256px)	Não especificado	AlexNet, AlexNetOWTB, GoogleNet, Overfeat, VGG	AlexNet, AlexNetOWTB, GoogleNet e VGG	AlexNetOWTB - 99,49; VGG - 99,53
30	Kaur et al. (2018)	PVD	Soja	4.775- Imagens RGB; Redimensionamento não especificado	Conversão para espaço de cor LAB e k-means	SVM	Color, GLCM, Gabor, 2DWT	SVM - ~ 90,0
31	Khan et al. (2018)	PVD e CASC-IFW	Citrino	Imagens RGB - (Exp 1: 2.688 imagens; Exp 2: 2.679 imagens; Exp 3, 4, 5, and 6: 1.200 imagens); tamanhos variados	Representação de espaço de cor: RGB, HSV e YCbCr	M-SVM (proposto); VGG16, Caffe AlexNet	Textura, Cor e LBP, incluindo Análise de Textura Fractal baseada em segmentação (SFTA)	M-SVM - 98,6
32	Hlaing et al. (2018)	PVD	7 Classes de Folha de Tomateiro	3.535- Imagens RGB	Limiar de cor -canal ver, preenchimento de região	SVM - Quadrática	Cor e Textura (SIFT)	SVM - 85,1 (10-fold)
33	Khan et al. (2018)	PVD e CASC-IFW		Imagens RGB - (Exp 1: 2.688; Exp 2: 2.679; Exp 3, 4, 5, e 6: 1.200); Tamanhos variados; Redimensionadas (224 x224) e (150x150px)	Fuzzy-C, K-means, Otsus e Segmentação da cor	M-SVM	LBP, Histograma de Cor (RGB, HSV, HSI)	Podridão Maça - 98,10; Crosia de Maça -97,30; Ferrugem de Maça -94,62; Saudável -98,0
34	Too et al. (2019)	PVD	14 Espécies de culturas	54.306- Imagens RGB; Tamanhos variados; Redimensionadas (224x224px)	Não especificado	ResNet 50, 101, 152 Camadas, VGG16, DenseNet e InceptionV4	CNN InceptionV4, VGG16, ResNet e DenseNet	DenseNet-99,75

35	Kamal et al. (2019)	PVD	55 Espécies de culturas	82.161- imagens RGB; Tamanhos variados; Redimensionadas (224 x224px) and (150x150px)	Não especificado	MobileNet Modificado, MobileNet Reduzido e MobileNet	CNN- MobileNet	MobileNet - 98,65; MobileNet Modif. - 97,5; MobileNet Red. - 98,34
36	Ahmed et al. (2019)	UCI	Airroz	120- Imagens RGB - aumentadas para 480 imagens	Não especificado	Regressão Logística (RL), KNN, DT (J48) e, Naïve Bayes (NB)	Não especificado	RL -70,0%, KNN -75,5%, DT- 95,0%, NB - 58,8%
37	Doh et al. (2019)	Kaggle	Variedade de citrinos	Imagens RGB; Redimensionadas (250x250px)	Clustering com K-means	SVM, ANN	Textura, cor, forma, morfológica	SVM- 93,12; ANN- 88,96
38	Sehgal, A., & Mathur, (2019)	Não especificado	Não especificado	Não especificado	Clustering com K-means	NN,SVM,RF,NB,DT	Cor e Textura	SVM- 72,92; RF- 71,88; NB- 70,57; DT-64,0
39	Hossain et al. (2019)	Arkansas e Reddit-plant	5 classes folhas com doenças	237- Imagens RGB	Conversão de RGB para LAB e uso de K-means; Morfologia	KNN, GLCM	GLCM, Color, Textura	KNN- 96,76
40	Jaisakthi (2019)	PVD	Folhas de Uva	5.675- Imagens RGB, Redimensionamento não especificado	GrabCut; Limiarização Global	SVM, RF, AdaBoost	GLCM	SVM- 91; RF- 83; AdaBoost 74,9
41	Abdu et al. (2020)	PVD	Batata	1.400 - Imagens RGB; Tamanhos variados; Redimensionadas (256 x 256px)	Representação de espaço de cores HSV e CIE LAB	Modelo proposto (C-SVM)	Vetor de coerência de cores (CCV) e LPB	C-SVM - 99,67; AlexNet - 93,86; ResNet50 - 92,22; NasnetLarge - 95,18
42	Karthik et al. (2020)	PVD	Tomate	95.999 - Imagens RGB de treinamento e 24001 de validação; Tamanhos variados	Não especificado	CNN Baseada em aprendizagem residual	CNN	CNN (5-fold) - 98,0
43	Basaviah & Anthony (2020)	PVD	5 classes	500- Imagens RGB; Redimensionadas (500 x 500px)	Não especificado	DT, RF	Histogramas de cores, Momentos Hu, Haralick e LBP	DT- 90,0; RF- 94,0
44	Rao et al. (2020)	PVD	38 classes	54.306- Imagens RGB; Redimensionadas (256 x 256px)	Conversão em espaço de cor e realce	Lógica Neuro-Fuzzy	GLCM, Gabor, Curvelet e Características de Momentos	Lógica Neuro-Fuzzy - 93,18
45	Kurmi et al. (2020)	PVD	3 classes: Batata, Pimenta e Tomate	Batata -1.000; Pimenta - 2.475; Tomate - e Banco 3- 15.112 imagens RGB	Conversão em espaço de cor LAB e usada técnica de crescimento da região	Multi-layer perceptron (MLP) e SVM	SIFT	Acurácia- 94.35
46	Gunarathna & Rathmayaa (2020)	Não especificado	10 Classes de folhas	22.930- Imagens RGB; Redimensionadas (256x256px)	Não especificado	SVM, NN, CNN	GLCM, LBP, CNN	SVM (GLCM)- 95,8; KNN- 91,4%; SVM (LBP)- 92,3; KNN- 90,1; CNN- 98,0
47	Gunarathna and Rathmayaa (2020)	Não especificado	10 Classes de folhas de tomate	22.930- Imagens RGB; Redimensionadas (256x256px)	Não especificado	CNN, VGG16, InceptionV3, MobileNet	CNN, VGG16, InceptionV3, MobileNet	CNN- 90,0; VGG16- 94,0; Inception V3- 95,0; MobileNet- 97,0
48	Radhakrishnan et al. (2020)	PVD	Airroz	60.000- Imagens RGB; Redimensionadas (256 x 256px)	Técnicas aprimoradas - Não especificadas no estudo	CNN and SVM	CNN-AlexNet	CNN- 95,8; CNN+SVM- 96,8
49	Sharma et al. (2020)	Kaggle	19 classes	20.000- Imagens RGB	Conversão RGB para HSV e K-means	RL, KNN, SVM e CNN	RL, KNN, SVM e CNN	RL- 66,4; KNN- 54,5; SVM- 53,4; CNN- 98,0

50	Kaur et al. (2021)	PVD	3 classes: Batata, Pimenta e Tomate	20.639- Imagens RGB	K-means	Classificador Melhorado	GLCM, Gabor, SHIFT	Classificador combinando- 95,66
51	Azim et al. (2021)	UCI	Arroz	Não especificada a quantidade de imagens e a resolução	Conversão RGB para formato HSV e limiar binário é aplicado encontrar as manchas	XGBoost, SVM com kernel RBF	Características de cor, forma (Momentos Hu), textura (GLCM) e LBP	XGBoost –86,58%; SVM com kernel RBF– 81,67%
52	Singh et al. (2022)	PVD	Maçã, milho, plantas de batata, tomate e arroz	37.315- Imagens RGB; Tamanhos variados	Conversão de cor RGB para cinza e RGB para espaço de cor LAB e k-mean	LeNet, ShuffleNet, AlexNet, EffNet e MobileNet,	HOG, GLCM e Momentos de cor	MobileNet - 96,10 (melhor)
53	Dhar et al. (2022)	Kaggle	10 Classes de folhas	Número de imagens RGB não especificado; Redimensionadas (200x150px)	Não especificada	SVM, KNN, Adaboost	Características globais (Gist) e locals (LBP)	SVM -80,0; KNN -73,4; Adaboost - 80,9
54	Singh et al. (2022)	PVD	38 classes	54.323- Imagens RGB; Redimensionadas (300 x 450px)	RGB convertida para cinza e para espaço LAB	LeNet, ShuffleNet, AlexNet, EffNet e MobileNet	Momentos da cor, HOG, GLCM	MobileNet (Otimização Bayesiana SVM)- 92,2; MobileNet (RF)- 96,1

Fonte: Autor (2023).

Shetty et al. (2022) e Joseph et al. (2022) conduziram estudos no campo da saúde, empregando um banco de dados público e descritores de extração de características, como Momentos Hu, Texturas Haralick e Histograma de cor. O primeiro focou no diagnóstico de lesões de pele utilizando 10.015 imagens RGB, enquanto o segundo abordou o câncer de mama com 7.909 imagens. No primeiro estudo, vários algoritmos de classificação foram implementados, resultando nas seguintes pontuações, RF (87,0%), DT (68,0%), LR (58,0%), LDA (57,0%), SVM (53,0%), KNN (48,0%), NB (36,0%), e CNN (95,18%) e no segundo estudo, em diferentes épocas de treinamento momento da rede neural do tipo CNN, foram obtidas as seguintes pontuações: 97,87% (40x), 97,60% (100x), 96,10% (200x) e 96,84% (400x). Portanto, os estudos reelaram melhor pontuação na acurácia com a combinação de diferentes descritores de extração de características.

A maioria das pesquisas tem empregado técnicas de aprendizado profundo, e os modelos demonstram desempenho superior para identificação e classificação de doenças em plantas, destacando os resultados de estudos conduzidos por diversos autores, incluindo Mohanty et al. (2016), Brahim et al. (2017), Amara et al. (2017), Too et al. (2019), Ferentinos (2018), Abdu et al. (2020), Khan et al. (2018), Kamal et al. (2019), Karthik et al. (2020). Uma tendência observada em todos esses estudos é que aqueles que utilizaram um banco de dados público tendem a alcançar desempenhos superiores, acima de 98% em termos de acurácia, em contraste com aqueles que optaram por bancos de dados específicos. Um exemplo notável que ilustra essa tendência pode ser observado no estudo de Sharif et al. (2018), no qual foram utilizados tanto bancos de dados públicos quanto específicos. Notavelmente, o mesmo modelo treinado apresentou resultados distintos, atingindo uma acurácia superior de cerca de 97% nos experimentos com dados públicos e relativamente inferior, com cerca de 90,4% de acurácia com os dados específicos, e uma acurácia ainda menor de cerca de 89% com a combinação dos tipos de banco de dados. Essa mesma tendência de menor acurácia ao usar bancos de dados específicos também foi observada em outros estudos, conduzidos por Kawasaki et al. (2015), Fujita et al. (2016), Lu et al. (2017), DeChant et al. (2017), Ramcharan et al. (2017), Picon et al. (2019), Selvaraj et al. (2019), Hu et al. (2019), Ma et al. (2018) e Zhang et al. (2019). Assim, as imagens que compõem os bancos de dados públicos frequentemente são

adquiridas em condições mais propícias, ou seja, o processo de captura das fotografias geralmente ocorre em ambientes de laboratório, ao contrário das imagens capturadas em situações reais de cultivo para análise.

Neste estudo, empregou-se um banco de dados específico, incorporando um procedimento de segmentação da folha que exclui a presença de elementos indesejados capazes de prejudicar a detecção e classificação de patologias. Além disso, as métricas de desempenho obtidas em experimentos semelhantes foram levadas em consideração para fins de comparação.

2.6 MAPEAMENTO DE PATOLOGIAS DE PLANTAS USANDO IMAGENS DIGITAIS

A tecnologia geoespacial tem desempenhado um papel crucial desde as primeiras tarefas de levantamento do estado de saúde das culturas até a gestão dos dados coletados. Estudos conduzidos por Anwer & Singh (2019), Prasetyo et al. (2017), Roberts et al. (2021), Singh et al. (2023), Nageswara Rao & Lakshmikantha (2020) e Dong et al. (2019) aplicaram técnicas de detecção por meio do sistema global de navegação por satélites (GNSS) para monitorar a saúde das plantas. Isso não apenas possibilita a localização precisa de patologias nas plantas, mas também aprimora a qualidade e a disponibilidade das informações geográficas, que podem ser acessadas em formato digital.

As fotografias digitais são amplamente utilizadas na colheita de amostras de doenças e de pragas para diagnóstico. Uma característica distinta das câmeras digitais modernas é a sua capacidade de fornecer não apenas imagens visuais, mas também informações geoespaciais, incluindo coordenadas GPS (*Global Positioning System*) e metadados denominados *Exchangeable Image File Format* (EXIF), incorporados nas imagens (Sabtu et al., 2018). Estas imagens georreferenciadas oferecem detalhes sobre o que foi capturado, quando e onde a foto foi tirada. Dessa forma, as imagens com coordenadas GPS podem ser facilmente integradas em

sistemas de gestão de informações, facilitando o rastreamento e o manejo eficaz de doenças e pragas agrícolas.

No estudo de Sabtu et al. (2018), foi demonstrado que os agricultores podem manter seus registros agrícolas e diários de campo diretamente em seus dispositivos por meio da tecnologia celular. Isso possibilita a inserção e o acesso instantâneo às informações no local. Além disso, a tecnologia GPS foi aplicada para registrar automaticamente a localização das doenças em plantas, mapeando assim a distribuição da incidência de pragas e doenças nas culturas. Conforme mencionado por Deleon et al. (2017), a tecnologia GPS e GIS (*Geographic Information System*) é utilizada para criar mapas georreferenciados de diferentes culturas e propriedades do solo. Esses mapas fornecem aos profissionais de campo e produtores um novo conjunto de ferramentas para comunicação e gestão. A diversidade crescente de mapas de produção atualmente disponíveis contribui significativamente para a melhoria do manejo das lavouras.

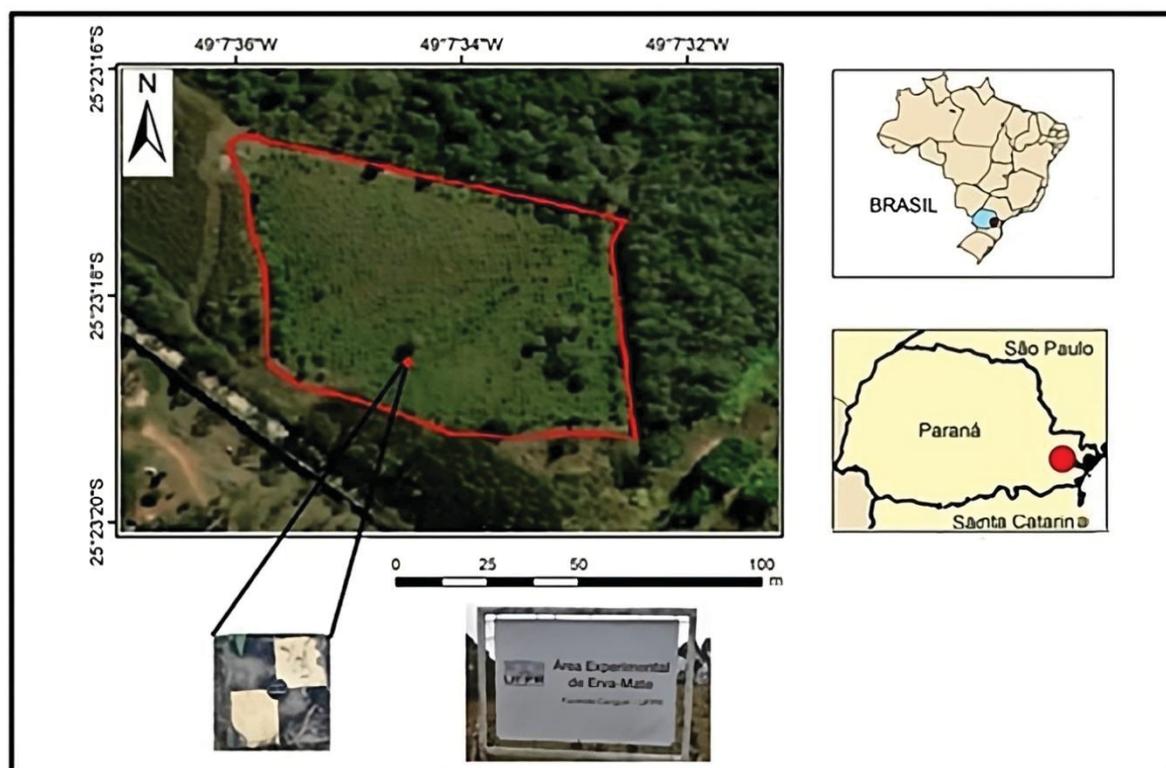
3 MATERIAIS E MÉTODOS

Este capítulo aborda os métodos utilizados na pesquisa e inclui uma breve apresentação da área de estudo, onde foram colhidas as amostras.

3.1 CARACTERIZAÇÃO DA ÁREA DE ESTUDO

O conjunto de dados de referência de imagens de doenças foliares de cultura de erva-mate foram coletadas no campo experimental da Fazenda Canguiri, da Universidade Federal de Paraná, Pinhais, Paraná. A área de estudo está localizada entre $49^{\circ}7' 36''$ W - $49^{\circ}7' 32''$ W e $25^{\circ}23' 20''$ S - $25^{\circ}23' 16''$ S, à uma altitude média de 914 m, precipitação média anual de 1.100 mm e uma temperatura média anual entre $19,9-26,5^{\circ}\text{C}$. Na Figura 5 é mostrado, em vermelho, o perímetro do campo ocupado por erva-mate.

FIGURA 5 - LOCALIZAÇÃO DA ÁREA DE ESTUDO



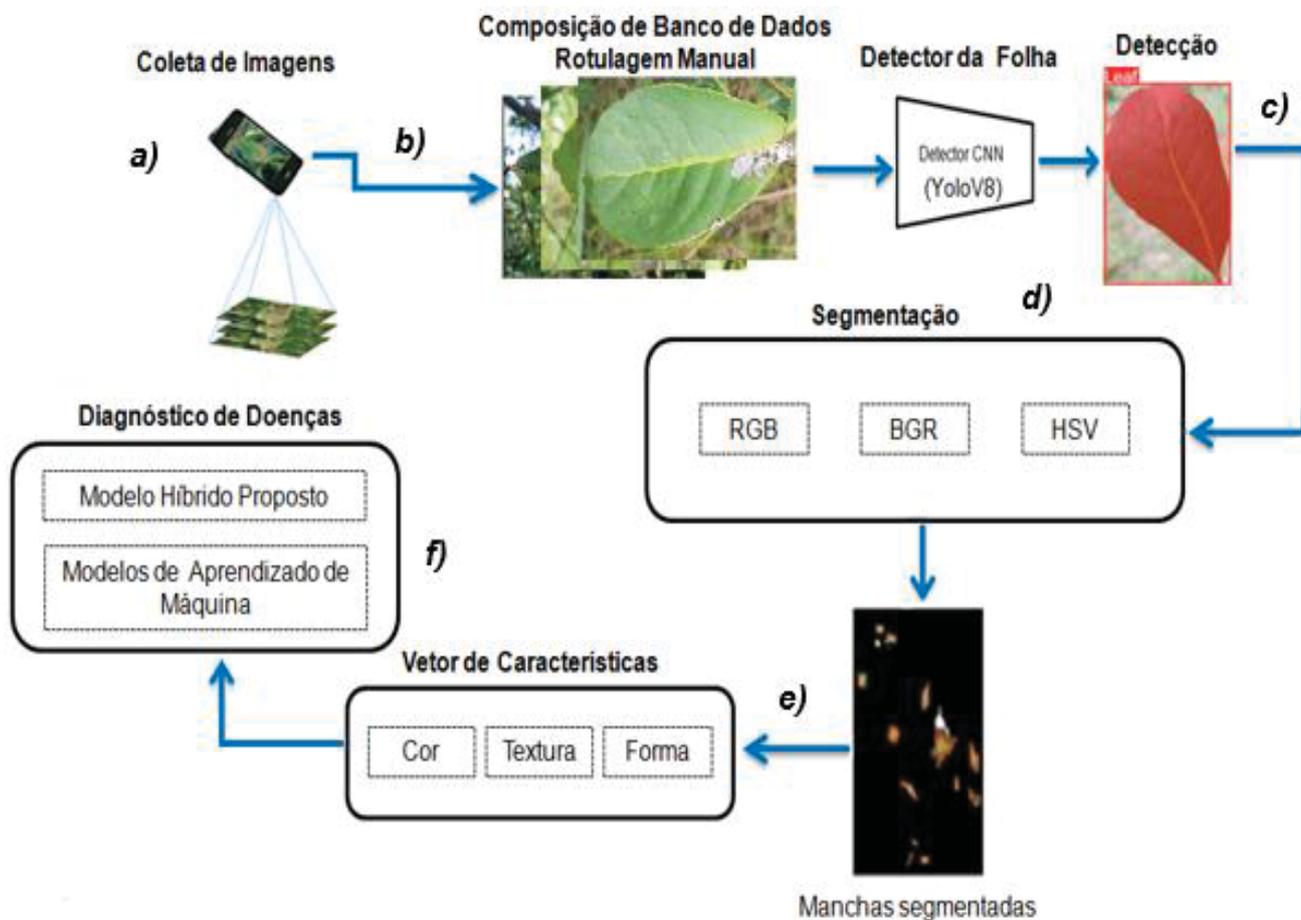
Autor (2023).

3.2 PROCEDIMENTO METODOLÓGICO

Diante das abordagens apresentadas na Tabela 4, na presente pesquisa, desenvolveu-se uma metodologia para detecção e classificação de doenças em plantas de erva-mate com base em sintomas visuais da folha. A abordagem adotada envolve a integração de características obtidas por métodos artesanais de extração e outras obtidas por aprendizagem profunda, especificamente o uso das CNNs, para a classificação. O desempenho dessa abordagem é considerado relativamente robusto, mas também merece uma investigação mais aprofundada e é importante destacar que, devido à sua arquitetura hierárquica, os modelos de aprendizado profundo têm a capacidade de aprender automaticamente recursos de alto nível a partir de dados brutos. Portanto, é fundamental reconhecer que o desempenho da CNN está diretamente relacionado à quantidade de amostras de treinamento rotuladas disponíveis.

O fluxograma metodológico é resumido na Figura 6 e compreende as seguintes etapas: *a)* coleta de imagens de folhas; *b)* pré-processamento; *c)* detecção; *d)* segmentação das folhas nas imagens; *e)* obtenção de parâmetros espectrais, espaciais e texturais e; *f)* classificação.

FIGURA 6 - FLUXOGRAMA METODOLÓGICO DA PESQUISA



Fonte: Autor (2023).

3.2.1 Coleta de Imagens e Composição do Banco de Dados

A disponibilidade de um banco de dados representativo é um dos maiores desafios nos projetos de desenvolvimento de sistemas de visão computacional que utilizam modelos de aprendizado profundo. Esta dificuldade se torna ainda maior quando se trata de resolver um problema específico, como o de detecção e classificação automática de patologias com base em imagens ópticas de folhas. Formar um banco de dados primário específico de treinamento suficiente utilizando é uma tarefa demorada e difícil. Este foi o primeiro passo da metodologia. O estudo foi desenvolvido utilizando um conjunto de dados primário específico, criado pelo autor.

Para isto, foram fotografadas grandes quantidades de imagens de folhas com fundo complexo, ou seja, folhas obtidas em ambiente de campo real de cultivo. O banco de imagens criado atende o desafio de desenvolver um método de segmentação de folhas de cultura de erva-mate em condições reais de campo de cultivo.

Conforme descrito no fluxograma (Figura 6), o processo começa com a aquisição de imagens. As imagens foram fotografadas entre Julho de 2022 a Agosto de 2023 utilizando dispositivo móvel - celular Samsung Galaxy A50, modelo da câmera SM-A505GT, com uma lente de 25 megapixels e resolução de 5774x4330 pixels e 1920x1080 pixels. Neste período foi possível coletar as imagens abrangendo diferentes estações do ano, permitindo compor dados que inclui informações detalhadas sobre as condições específicas em que foram coletadas. Foram coletadas 4.660 imagens RGB, armazenadas no formato JPG, com resolução de 3024x3024 pixels (resolução original da imagem capturada no campo). O conjunto de dados foi composto por 4 classes de imagens: (a) *folhas saudáveis*, (b) *patologia pinta-preta*, (c) *patologia fumagina*, e (d) *patologia requeima*.

Na Figura 7 mostra um exemplo da planta de erva-mate e o ambiente na qual as folhas foram fotografadas.

FIGURA 7 - EXEMPLO DA PLANTA DE ERVA-MATE



Fonte: Autor (2023).

As folhas eram centralizadas aproximadamente na moldura a ser fotografadas, numa proporção 1:1 e cada classe e tipo de doença era distinta.

Portanto, foram coletadas 1.165 imagens de folhas para cada categoria, assegurando uma distribuição equitativa de amostras e uma representação adequada de todas as classes. Todas as imagens em este conjunto de dados foram obtidas em condições relativamente boas, com luz solar difusa para reduzir a criação da sombra nas imagens fotografadas.

No processo de coleta de dados, foi adotada uma abordagem de amostragem não probabilística. Ao realizar a documentação das patologias no campo experimental, quase todo o campo foi percorrido sistematicamente, examinando cada planta para verificar a presença de diferentes tipos de doenças. Essa metodologia de amostragem foi guiada pela observação direta, sem seguir um método probabilístico. As amostras foram selecionadas com base na disponibilidade e na identificação visual das patologias, garantindo a representação abrangente de todas as classes de interesse. A escolha de capturar 1.165 imagens para cada categoria reflete a busca por uma representação equitativa das classes, contribuindo para um banco de dados robusto e abrangente para análise e desenvolvimento do modelo.

No conjunto de dados, foram realizados dois tipos de rotulagem. Inicialmente, as imagens foram categorizadas com base em seu conteúdo visual, considerando diversas características distintivas entre as classes. Esse processo de rotulagem foi conduzido no campo experimental de plantio de erva-mate, com a colaboração de uma especialista, técnica agrônoma responsável pelo campo experimental. Essa profissional desempenhou um papel crucial na diferenciação das características de cada patologia, conforme descrito por Grigoletti Júnior e Auer (2013), Kimati et al. (2005), e pelo Embrapa Florestas (Colombo, PR) no "Manual de Identificação de Pragas e Doenças da Erva-Mate" (*Ilex paraguariensis* St. Hil. Colombo: Embrapa Florestas, 2000, 23p, Embrapa Florestas. Documentos, 44).

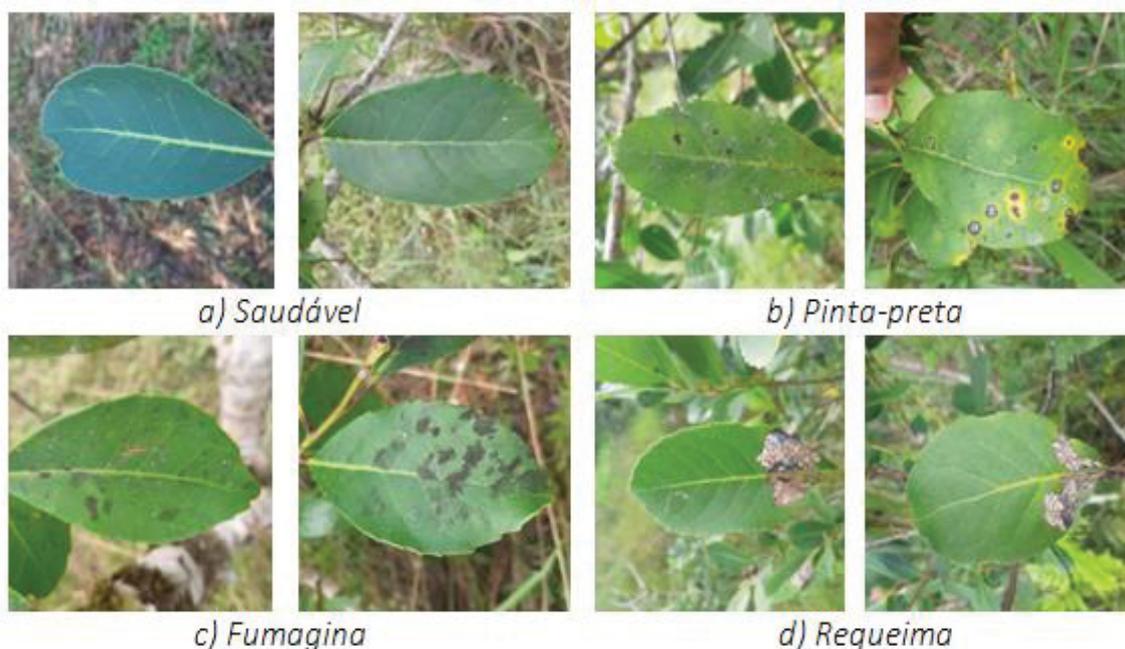
O quadro 1 apresenta as patologias identificadas, e a caracterização de cada uma permitiu a rotulagem visual apropriada. Por exemplo, a rotulagem das classes foi realizada da seguinte maneira: (i) folhas sem patologia (Saudável) considerou-se todas as folhas sem qualquer sintoma visual ou com 98% e 99% da parte saudável, sendo classificadas como folhas saudáveis; (ii) a patologia Pinta-preta, causada pelo

fungo *Cylindrocladium spathulatum*, apresenta lesões foliares arredondadas, por vezes concêntricas, no interior ou nos bordos do limbo, principalmente em folhas adultas. Essas manchas podem atingir até 2 cm de diâmetro, aumentando em tamanho e número, cobrindo grande parte da folha e resultando na sua queda prematura. Na face abaxial da folha, em condições de umidade constante, observa-se uma frutificação esbranquiçada, sinal distintivo do patógeno; (iii) a patologia Fumagina, causada pelos fungos *Meliola sp.* e *Capnodium sp.*, apresenta um sintoma visual característico, em que os fungos associados formam uma crosta espessa e escura sobre folhas e ramos, aproveitando-se das secreções de insetos como formigas, cochonilhas e pulgões, que geralmente estão associados a esses sintomas. O ataque pode ocorrer em parte ou em toda a planta; e (iv) a patologia Requeima, causada pelo fungo *Phytophthora infestans*, apresenta sintomas variados nas folhas. Estas exibem tamanho reduzido e formato irregular, com coloração que varia de verde-claro a escuro. Um halo encharcado é visível, delineando a lesão do tecido sadio. Em alguns casos, observa-se um crescimento esbranquiçado no interior da folha, correspondente às estruturas reprodutivas do fungo. Em situações distintas, o crescimento das lesões é interrompido, e o tecido torna-se quebradiço.

O processo de rotulagem manual foi crucial para fornecer os rótulos corretos nos dados de treinamento e validação, permitindo o desenvolvimento de um modelo de classificação preciso. A seleção dessas patologias como objetos de investigação fundamenta-se na sua maior ocorrência no contexto analisado, sendo patologias frequentes em plantas de erva-mate. Além disso, essas patologias foram escolhidas por serem facilmente identificáveis no campo experimental de erva-mate, o que facilita a análise e a aplicação de métodos de detecção. A prevalência dessas patologias torna-as representativas dos desafios enfrentados na agricultura da erva-mate, tornando os resultados do estudo mais aplicáveis e generalizáveis para cenários práticos. Essa escolha estratégica permite uma abordagem mais focalizada, concentrando-se em patologias de maior relevância e impacto econômico, ao mesmo tempo em que simplifica o processo de classificação e validação das imagens no desenvolvimento do modelo de detecção. Dessa forma, a pesquisa busca abordar questões práticas e prioritárias no contexto da agricultura de erva-mate, maximizando a utilidade e a aplicabilidade dos resultados obtidos.

A Figura 8 ilustra a composição do banco de dados, oferecendo uma visão geral da amostra de folhas da cultura de erva-mate: (a) *Folhas Saudáveis*, (b) *Patologia Pinta-preta*, (c) *Patologia Fumagina* e, (d) *Patologia Requeima*.

FIGURA 8 - DIFERENTES CATEGORIAS DO CONJUNTO DE DADOS



Fonte: Autor (2023)

Após a composição de banco de dados, as imagens das quatro classes foram rotuladas manualmente (vide Figura 9), delimitando em torno da imagem da folha, de forma a fornecer informações inerentes às imagens para o algoritmo de detecção e segmentação de instâncias. Isso consiste em atribuir rótulos às imagens para o treinamento do modelo de aprendizado profundo selecionado YOLOv8. A rotulação foi realizada manualmente utilizando a ferramenta de anotação de visão computacional de código aberto denominado CVAT (*Computer Vision Annotation Tool*) desenvolvida pela Intel. Nesse processo, os arquivos de anotação foram salvos em formato PNG e posteriormente convertidos para o formato TXT para usar o algoritmo.

O modelo YOLOv8 requer algumas informações inerentes às imagens, como as coordenadas do contorno do objeto a detectar, ponto médio (x, y), a largura (w) e a altura (h) da caixa delimitadora, além da definição da classe do objeto a ser

detectado. Realizou-se anotação com base no uso de polígono, delimitando as folhas em primeiro plano para sua detecção e segmentação. Este processo, apesar de ser uma tarefa laboriosa por ser feita manualmente é consideravelmente acurada. Ainda assim, o sucesso desta anotação é altamente dependente da participação de um ser humano treinado para evitar erros. A Figura 9, mostra as coordenadas de contorno da folha e caixa delimitadora (x_1, y_1) e (x_2, y_2) , centro da caixa delimitadora (x, y) , a altura (h) e largura (w) da caixa delimitadora.

FIGURA 9 - FIGURA 9 - COORDENADAS DE CONTORNOS DA FOLHA E CAIXA DELIMITADORA



Fonte: Autor (2023).

3.2.2 Detecção

Na tarefa de detecção, o objetivo era separar as folhas do plano de fundo, o qual pode conter diversas informações indesejadas. A segmentação foi realizada nas folhas relativamente maiores da imagem, previamente detectadas pelo algoritmo proposto. Cada folha foi delimitada por uma caixa delimitadora prevista, com uma probabilidade maior de que a janela da imagem contenha a folha. Nesta fase de detecção foi usada uma rede convolucional YOLO, do que deriva seu nome (*YOLO – you only look once*). Este tipo de detector se caracteriza pela sua rapidez, pois varrem a imagem apenas uma vez, reduzindo o tempo de processamento. No detector YOLO, o mapa de características é dividido em células e para cada célula, são treinadas camadas totalmente conectadas para detectar os objetos de interesse. O algoritmo YOLO trata o problema de detecção de objetos como um problema de regressão e o

resultado é a delimitação do local onde o objeto foi detectado (caixa delimitadora) e um valor de pontuação que descreve o grau de sucesso da detecção de cada caixa é encontrada, fixando um limiar torna-se assim possível descartar detecções “fracas” ou para as que existe pouca evidência na imagem (Redmon e Farhadi, 2018; Cheng, 2020).

A Figura 10 apresenta a estrutura simplificada de YOLOv8 padrão proposto na pesquisa, primeiro a folha da erva-mate na imagem é localizada e em seguida a segmentação de instância da folha detectada. O YOLOv8 foi desenvolvido pela *Ultralytics*⁴, que também criou o modelo YOLOv5, incluindo várias alterações e melhorias na experiência de arquitetura e desenvolvedor em relação ao YOLOv5.

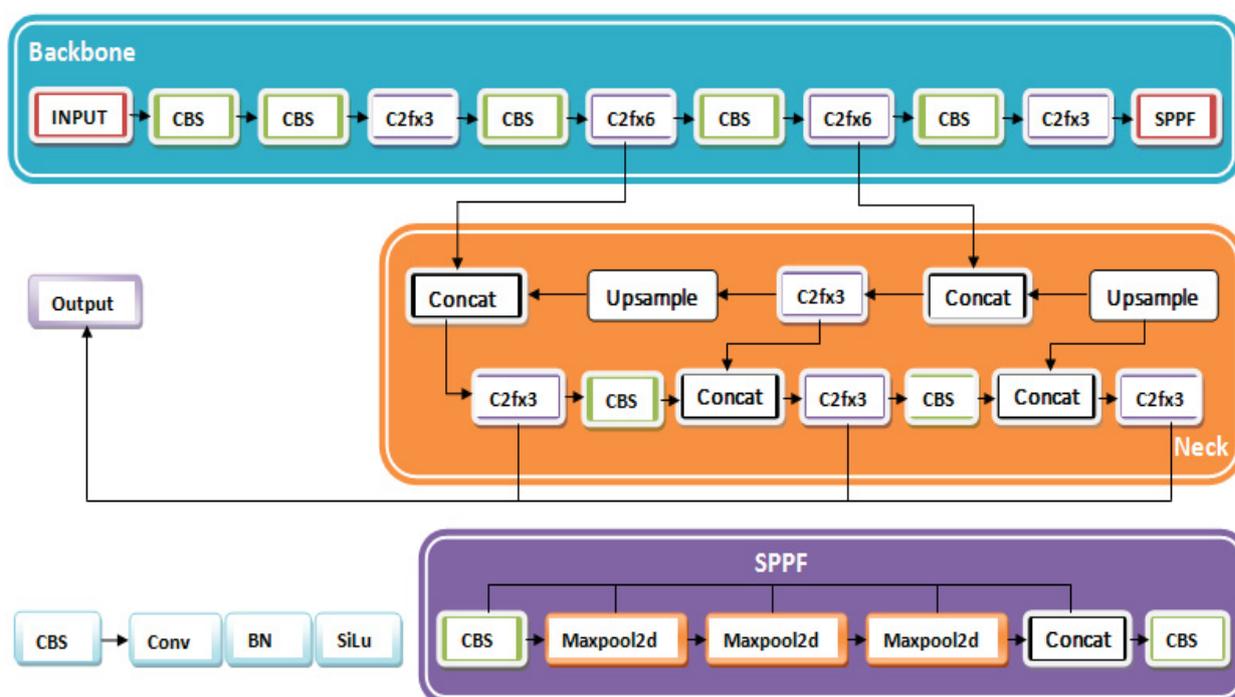
No presente trabalho, selecionou-se a variante YOLOv8n. YOLOv8n é uma estrutura de parâmetros leve derivada do algoritmo YOLOv8. Ele compreende a parte de *Backbone*, uma parte denominada *Neck* e a parte de saída de previsões. Na *Backbone* utiliza operações convolucionais para extrair características de várias escalas de imagens coloridas RGB. Enquanto isso, o papel da *Neck* é mesclar as características extraídas pela rede do Backbone e a parte de saídas serve para prever categorias e detectar objetos em diferentes tamanhos. Utiliza *Feature Pyramid Networks* (FPN) para agregar características em diferentes níveis, desde baixo nível até superior. A rede possui uma camada final que faz previsões sobre a categoria do objeto, enquanto três conjuntos de detectores (com tamanhos variados) são utilizados para analisar a imagem em diferentes escalas, permitindo uma detecção mais abrangente de objetos de diferentes tamanhos. YOLOv8 tem vários casos de uso em tarefas de detecção de objetos e classificação de imagens, entre veículos autônomos, vigilância, varejo, imagens médicas, robótica, agricultura, etc.

A parte de backbone do YOLOv8 é basicamente a mesma do YOLOv5, com uma mudança no modelo C3, que na YOLOv8 foi substituído pelo módulo C2f inspirado na ideia do CSP (*Cross Stage Partial Networks*) (Jocher et al., 2023; Lou et al., 2023). O módulo C2f é uma modificação inspirado com a ideia baseada em ELAN (*Extraordinary Linear Attention Network*) presente no YOLOv7 (Wang et al., 2023).

⁴ <https://github.com/ultralytics/ultralytics/issues/189>.

Essa adaptação combina elementos dos módulos C3 e ELAN para formar o C2f, com o objetivo de obter informações de fluxo de gradiente mais ricas. No final do backbone, o módulo SPPF (*Spatial Pyramid Pooling Fusion*), é mantido. Ele utiliza três operações Maxpooling de tamanho 5x5 em série, e as saídas dessas operações são concatenadas. Isso é feito para garantir a precisão na detecção de objetos em diferentes escalas. Por outro lado, há uma ênfase na leveza do modelo YOLOv8, buscando manter um peso reduzido enquanto ainda preserva a capacidade de detectar objetos em várias escalas (Jocher et al., 2023; Lou et al., 2023).

FIGURA 10 - ESTRUTURA DA ARQUITETURA DA REDE YOLOV8



Fonte: Autor (2023).

Em geral, o YOLOv8 adapta e combina elementos de versões anteriores (v5 e v7), introduzindo modificações específicas no backbone para melhorar a representação de características e garantir a leveza do modelo.

No módulo *Neck* na arquitetura YOLOv8 utiliza o método de fusão de características PAN-FPN (*Path Aggregation Network with Feature Pyramid Network*). O PAN-FPN é uma extensão FPN e visa resolver limitações relacionadas à fusão de

características em diferentes níveis de uma pirâmide. A ideia central é realizar uma agregação de caminho, onde as informações são propagadas em direções ascendentes e descendentes, ajudando a melhorar a comunicação entre camadas de diferentes resoluções. Ainda são utilizados dois módulos de upsampling, com vista a preservar os detalhes finos e a melhorar a precisão das previsões e vários módulos C2f, juntamente com uma estrutura *Neck* desacoplada. A ideia de desacoplar o módulo *Neck* introduzida no YOLOv8, envolve a combinação de caixas de confiança e regressão para melhorar a precisão da detecção, ou seja, permite processar de forma independente tarefas de objetividade, classificação e regressão. Por outro lado, essa abordagem visa aprimorar a representação de características e a fusão de informações para a detecção de objetos em diferentes escalas. O SPPF (*Spatial Pyramid Pooling Fusion*) é composto por três níveis de Maxpool e dois CBS (*Conv-BN-SiLU*). O objetivo do SPPF é fornecer uma representação em múltiplas escalas dos mapas de características de entrada. O CBS é composto por funções de convolução, normalização de lote e funções de ativação SiLU (*Sigmoid Linear Unit*) e utiliza a função softmax para as probabilidades das classes, representando as probabilidades dos objetos pertencentes a cada classe possível. Essa combinação de operações é usada para processar as características extraídas no módulo *Neck*, contribuindo para a eficácia do modelo (Jocher et al., 2023; Lou et al., 2023).

De acordo com Lou et al. (2023), YOLOv8 usa modelo sem âncoras, o que significa que não depende de caixas de âncoras pré-definidas para detectar objetos. Ele calcula o grau de alinhamento do nível de âncora para cada instância usando a Equação (1):

$$t = s^\alpha * u^\beta \quad (1)$$

onde s é a pontuação de classificação, u é o valor de IoU e α e β são pesos de hiperparâmetros. O modelo seleciona âncoras com o valor máximo (t) em cada instância como amostras positivas e seleciona as outras âncoras como amostras negativas e, em seguida, treina por meio da função de perda. Após as melhorias acima, o YOLOv8 é 1% mais preciso que o YOLOv5 (Lou et al., 2023), tornando-o o detector mais preciso. O YOLOv8 pode suportar todas as versões do YOLO e pode

alternar entre diferentes versões. Também pode rodar em várias plataformas de hardware (CPU-GPU), dando-lhe grande flexibilidade.

O YOLO usa uma função de perda específica chamada *Cross-Entropy Loss* (em português Perda de Entropia Cruzada Binária), também conhecida como *log loss*. A *Cross-Entropy Loss* é frequentemente utilizada em tarefas de classificação e detecção de objetos, em vez de MSE (soma do erro ao quadrado), que é mais comum em problemas de regressão. Essa função de perda compara as previsões do modelo com as verdadeiras classes dos objetos, penalizando mais fortemente as previsões incorretas. Para a detecção de objetos no YOLO, a função de perda inclui termos relacionados à confiança (probabilidade de objeto estar presente), coordenadas da caixa delimitadora e classes dos objetos. A ideia é penalizar as previsões que estão longe da verdadeira localização e classe do objeto

A perda de coordenadas avalia a diferença entre as coordenadas preditas e as coordenadas verdadeiras das caixas delimitadoras. Essa medida quantifica o quão precisa é a detecção da localização dos objetos na imagem. A perda de confiança calcula a diferença entre as pontuações de confiança preditas e as pontuações verdadeiras. Essa perda permite avaliar o quão confiável é a detecção do objeto, ou seja, a probabilidade de que a caixa delimitadora proposta contenha realmente um objeto. A perda de classe mede a discrepância entre as probabilidades de classe preditas e as probabilidades verdadeiras. Essa perda é utilizada para avaliar a precisão da classificação do objeto, ou seja, a probabilidade de que a classe atribuída à caixa delimitadora esteja correta.

Portanto, a função de perda do YOLO combina essas três perdas em uma única medida de erro, permitindo que o modelo seja treinado de forma a minimizar a diferença entre as previsões e os valores verdadeiros. Essa abordagem de perda múltipla é fundamental para o treinamento eficaz do YOLO e para melhorar a precisão da detecção e classificação de objetos nas imagens (Equação 2):

$$\text{loss} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} [(x_i - \hat{x})^2 + (y_i - \hat{y}_i)^2] \quad (2)$$

$$\begin{aligned}
& + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} \left[(\sqrt{w_i} - \sqrt{\widehat{w}_i})^2 + \sqrt{h_i} - \sqrt{\widehat{h}_i} \right]^2 \\
& + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} (C_i - \widehat{C}_i)^2 + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{noobj}} (C_i - \widehat{C}_i)^2 \\
& + \sum_{i=0}^{S^2} 1_i^{\text{obj}} \sum_{c \in \text{classes}} (p(c) - \widehat{p}_i(c))^2
\end{aligned}$$

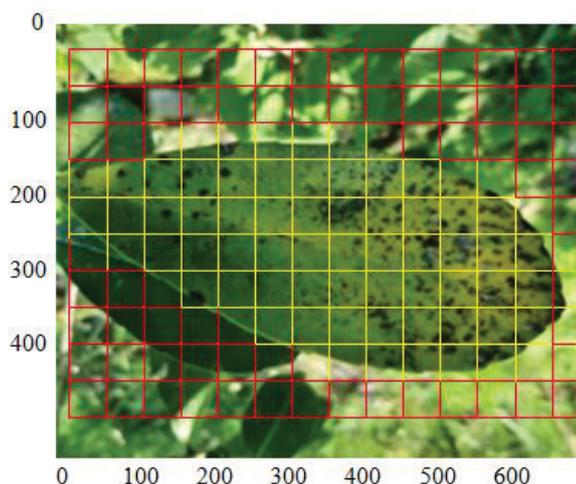
onde λ_{coord} , λ_{noobj} , λ_{classes} são hiperparâmetros que controlam a contribuição de cada componente de perda. O objetivo da função é minimizar a diferença entre as previsões do modelo e as verdades.

Para o estudo, o YOLOv8 foi ajustado e treinado utilizando diferentes imagens de folhas de erva-mate. O algoritmo emprega uma abordagem na qual uma única rede neural é aplicada à imagem completa. Posteriormente, a imagem é subdividida em regiões, e o algoritmo gera previsões de caixas delimitadoras candidatas, acompanhadas das respectivas probabilidades de confiança associadas a cada uma dessas caixas (Bochkovskiy, et al., 2020). As caixas delimitadoras passam por um refinamento adicional por meio de uma camada de regressão. Essa camada ajusta os valores das coordenadas da caixa delimitadora para otimizar sua correspondência com o objeto alvo. A regressão logística é empregada, e uma função de perda para prever as pontuações de confiança das caixas delimitadoras conterem objetos e ajustar os pesos do modelo durante o treinamento. No contexto da predição de classes, aplicado, por exemplo, a folhas de erva-mate, o modelo realiza previsões das probabilidades das classes dos objetos contidos no conjunto de dados para cada caixa delimitadora candidata.

Na etapa de pós-processamento, todas as caixas delimitadoras geradas são agrupadas e filtradas de acordo com as probabilidades de confiança associadas a cada caixa, utilizando limiares definidos pelo usuário. Em seguida, as caixas delimitadoras sobrepostas são unificadas por meio da função *Non-Maximum Suppression* (NMS) para produzir as caixas delimitadoras finais. O NMS é usado para remover redundâncias, mantendo apenas as caixas delimitadoras mais confiantes e eliminando as sobreposições indesejadas. A detecção sem âncora reduz o número de previsões de caixa, o que acelera a supressão não máxima (NMS), uma etapa

complicada de pós-processamento que filtra as detecções de candidatos após a inferência. Na Figura 11 é mostrada exemplo de uma célula da grade que contém a folha. Cada célula da grade é responsável por criar outras caixas delimitadoras possíveis com sua pontuação de confiança e probabilidades de detectar o objeto.

FIGURA 11 - REPRESENTAÇÃO DA CÉLULA DA GRADE QUE CONTÉM O OBJETO



Fonte: Autor (2023).

A Intersecção sobre a União (IoU) é uma métrica utilizada para medir a sobreposição entre caixas delimitadoras em detecção de objetos (Equação 3). Ela é calculada dividindo a área da intersecção entre as caixas pela área da união entre elas. O resultado é um valor entre 0 e 1, indicando a proporção da sobreposição. No caso do nosso estudo, o IoU é empregado para avaliar a precisão da localização das delimitadoras. Quando uma célula da grade contém um objeto, a pontuação de confiança atribuída a essa caixa delimitadora é igual ao valor do IoU. Se não houver nenhum objeto presente na célula, a pontuação de confiança é definida como zero. Quando a pontuação de IoU é igual a 1, significa que a área de intersecção entre a caixa delimitadora prevista e a caixa real do objeto é igual à área da união dessas duas caixas. Em outras palavras, não há diferença entre a caixa delimitadora prevista e a real, indicando uma correspondência perfeita e uma localização precisa do objeto. Essa é uma condição ideal e desejada em tarefas de detecção de objetos, indicando que o modelo fez uma previsão precisa em relação à posição e tamanho do objeto. Por outro lado, valores de IoU menores indicam uma sobreposição menor ou nenhuma

sobreposição entre as caixas delimitadoras. Através da (Equação 3) descrito no trabalho de Redmon et al. (2016) é possível avaliar a qualidade das previsões do modelo em relação à localização dos objetos na imagem.

$$\text{IoU} = \frac{\text{Área de Interseção}}{\text{Área da União}} \quad (3)$$

A pontuação de confiança indica o grau de certeza do modelo de que a caixa contém um objeto, ou seja, o quão preciso ele considera a caixa que prevê. A pontuação de confiança é calculada pela (Equação 4), descrito no trabalho de Redmon et al. (2016):

$$\text{Confiança} = \text{Pr}(\text{Objecto}) * \text{IoU}_{\text{previsto}}^{\text{verdadeiro}} \quad (4)$$

onde, $\text{Pr}(\text{Object})$ é a probabilidade de existir um objeto na caixa delimitadora predita pelo modelo e; $\text{IoU}_{\text{previsto}}^{\text{verdadeiro}}$ é a interseção sobre união (IoU) entre a caixa delimitadora predita e a caixa delimitadora verdadeira. Olhando para Figura 3, é possível se ter ideia de como o modelo calcula as pontuações de confiança específicas da classe para cada caixa. Essas probabilidades condicionais de classe são multiplicadas por previsões de confiança de caixa individual que fornecem pontuações de confiança específicas de classe para cada caixa, conforme mostrado na (Equação 5), descrito no trabalho de Ünver and Ayan (2019):

$$\text{Prob.} \left(\frac{\text{Classe}}{\text{Objeto}} \right) * \text{IoU}_{\text{previsto}}^{\text{verdadeiro}} = \text{Prob.} (\text{Classe}) * \text{IoU}_{\text{previsto}}^{\text{verdadeiro}} \quad (5)$$

Diversas pesquisas investigaram as funções de perda em uma variedade de algoritmos de detecção de objetos de última geração, como YOLOv3, SSD, Faster R-CNN, entre outros.

Os experimentos de detecção e segmentação da folha de erva-mate foram realizados no ambiente do Windows 11, com sistema operacional de 8 GB de memória RAM, configurado para utilizar a GPU CUDA, especificamente uma NVIDIA T600 com 4096 MiB de memória - implementado em Python 3.11.4, usando PyTorch 2.0.1 com suporte para CUDA 11.8.

O conjunto de dados de referência foi dividido em conjuntos de treinamento (70%) e validação (20%). Outro conjunto de dados selecionado independentemente dos conjuntos de treinamento e validação foi utilizado para o teste do modelo, correspondente a 10%. Essa divisão é importante para avaliar o desempenho dos modelos em dados não vistos anteriormente e garantir a generalização dos resultados.

Para treinar o modelo, certos hiperparâmetros foram configurados. Foi utilizado otimizador *Stochastic Gradient Descent* (SGD), taxa de aprendizado: 0.01, valor de *momentum*: 0.9, *batch size*: 8, *workers*: 2, número de épocas: 200 e; os pontos de verificação (*checkpoints*) salvos a cada 10 épocas. Antes de iniciar o treinamento do modelo, as imagens de entradas foram redimensionadas para a resolução de 640 x 640 pixels. O redimensionamento é aplicado por diversas razões, principalmente relacionadas à eficiência computacional e desempenho do modelo. Ao padronizar as dimensões das imagens de entrada, facilita-se o processamento paralelo e otimização durante o treinamento, resultando em tempos de treinamento mais rápidos e exigindo menos recursos computacionais. Além disso, a uniformização das dimensões simplifica a implementação de arquiteturas de redes neurais, pois os tensores de entrada têm tamanhos consistentes. Isso contribui para um treinamento mais estável e evita desafios associados a variações nas dimensões das imagens.

Para realizar uma avaliação quantitativa do desempenho do método de segmentação proposto, YOLOv8, foi conduzido um experimento comparativo para segmentação de folhas de cultura de erva-mate com fundo complexo. Esse método foi comparado com outros métodos de segmentação, incluindo o método tradicional GrabCut, o método de aprendizagem profunda tradicional U-Net e o método de aprendizagem profundo de dois estágios, Mask R-CNN. Os métodos são descritos a seguir:

a) *GrabCut*: GrabCut é um método de segmentação de imagens que se baseia em cortes de gráficos proposto por Rother et al. (2004). O processo se inicia com a definição de uma caixa delimitadora especificada pelo usuário ao redor do objeto a ser segmentado. O algoritmo estima a distribuição de cores do objeto alvo e do fundo

usando um modelo de mistura gaussiana. Isso é usado para construir um campo aleatório de Markov sobre os rótulos de pixel, com uma função de energia que prefere regiões conectadas com o mesmo rótulo, e executar uma otimização baseada em corte de gráfico para inferir seus valores. Como esta estimativa provavelmente será mais precisa do que a original, retirada da caixa delimitadora, este procedimento de duas etapas é repetido até a convergência desejada.

O usuário tem a possibilidade de refinar ainda mais as estimativas, corrigindo regiões mal classificadas e executando novamente a otimização. Além disso, o método implementa correções nos resultados visando preservar as arestas da segmentação. Existem várias implementações de código aberto disponíveis, incluindo OpenCV (a partir da versão 2.1).

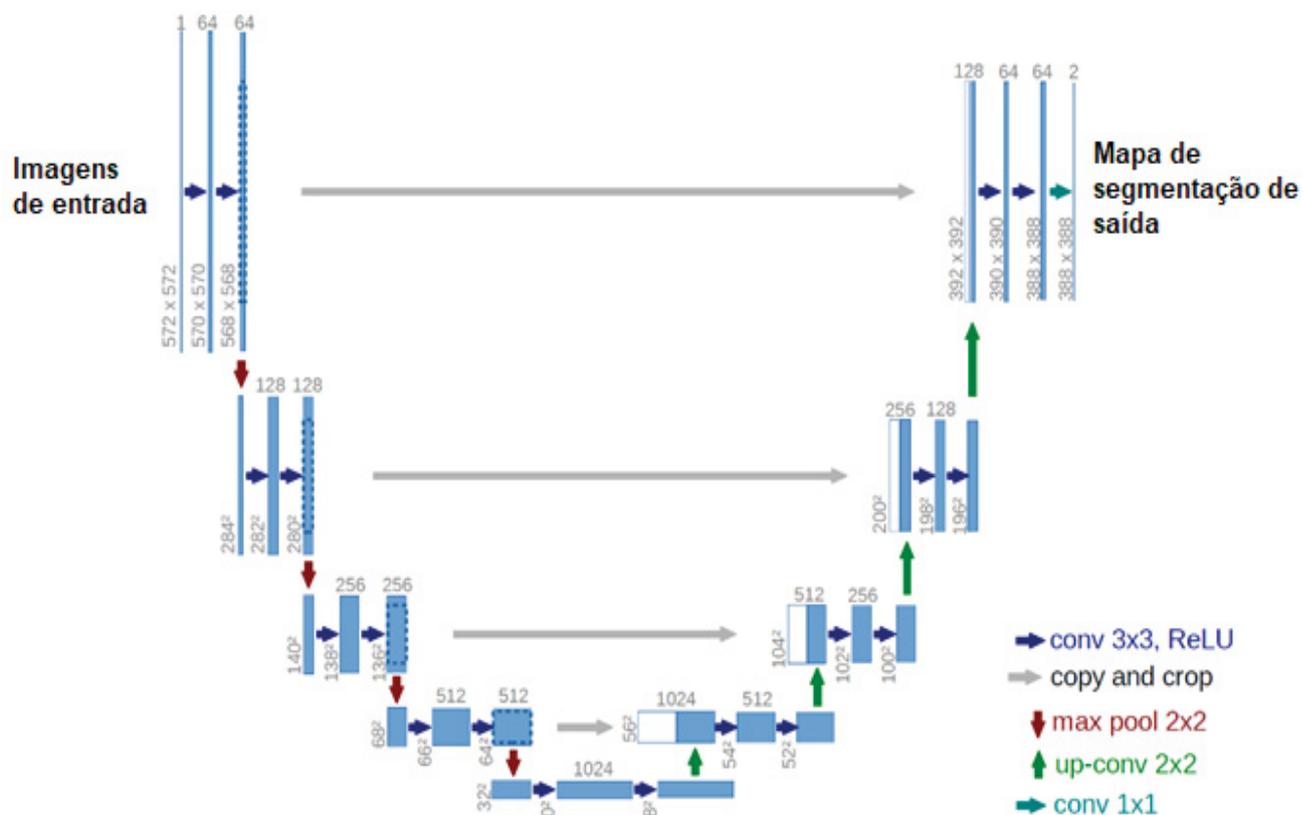
b) U-Net: A U-Net representa uma arquitetura de rede convolucional concebida por Ronneberger et al. (2015), projetada para realizar segmentação de imagens de forma rápida e precisa, notadamente aplicada à segmentação de imagens biomédicas. Essa abordagem demonstrou seu desempenho superior ao superar a convolucional de janela deslizante no desafio ISBI, focado na segmentação de estruturas neuronais em pilhas de microscopia eletrônica. Além disso, obteve êxito ao vencer desafios relevantes, como a detecção automatizada de cáries em radiografias Bitewing no ISBI 2015 e o rastreamento celular no mesmo ano, destacando-se nas categorias mais desafiadoras de microscopia de luz transmitida.

Em outras pesquisas, esse algoritmo foi implementado em diversas áreas, com destaque para a segmentação de objetos, conforme evidenciado pelos estudos de Patidar e Vianny (2020), Jia et al. (2022) e Chen et al. (2021). Essas investigações adicionais indicam a adaptabilidade e a eficácia da U-Net em contextos variados, solidificando sua relevância em diversas aplicações para além da segmentação biomédica.

A arquitetura do modelo é bastante simples: um codificador (para redução da resolução) e um decodificador (para aumento da resolução) com conexões de salto. Como mostra a Figura 12, ele tem o formato da letra U, daí o nome U-Net. As setas cinza indicam as conexões de salto que concatenam o mapa de recursos do

codificador com o decodificador, o que ajuda no fluxo reverso de gradientes para melhorar o treinamento. Portanto, com a compreensão da segmentação semântica e da arquitetura U-Net, implementou-se a arquitetura U-Net com TensorFlow e Keras no *notebook Colab* para a segmentação das folhas de erva-mate.

FIGURA 12 - ARQUITETURA DO MODELO U-NET



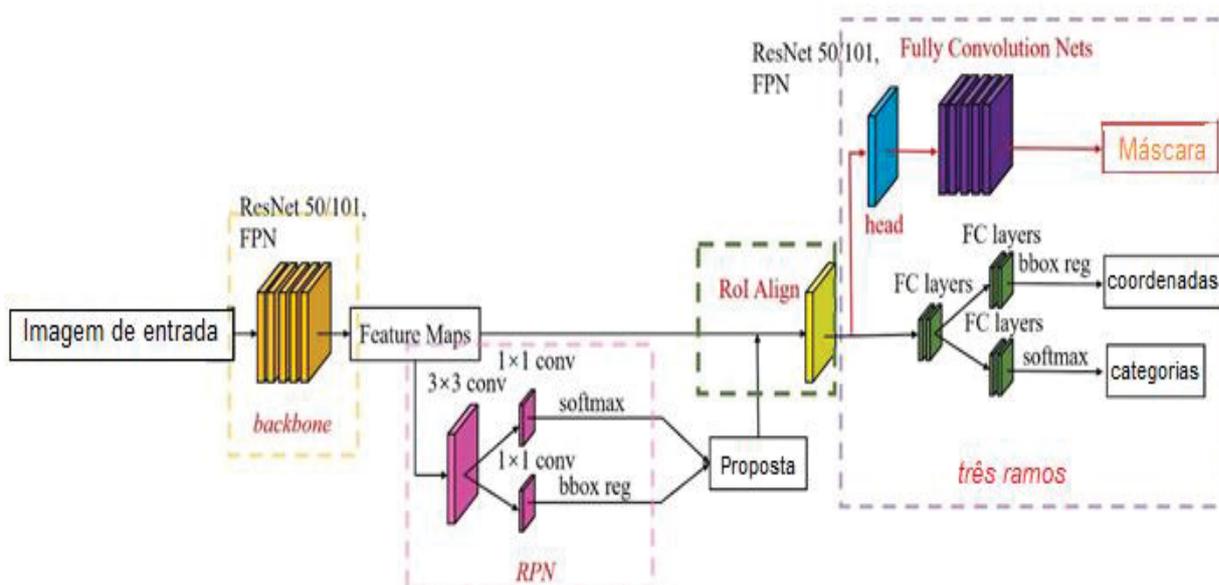
Fonte: Ronneberger et al. (2015).

c) *Mask R-CNN*: Mask R-CNN é um modelo de aprendizado profundo que combina detecção de objetos e segmentação de instâncias, proposto por He et al. (2017), representando uma extensão da arquitetura Faster R-CNN. Sua principal inovação reside na habilidade de realizar segmentação de instâncias em nível de pixel, simultaneamente à detecção de objetos. Isto é conseguido através da adição de um ramo extra de "mask head", responsável por gerar máscaras de segmentação precisas para cada objeto identificado. Essa abordagem proporciona limites refinados em nível de pixel, resultando em uma segmentação de instâncias precisa e detalhada.

Dois aprimoramentos fundamentais incorporados ao Mask R-CNN são o ROIAlign e a *Feature Pyramid Network* (FPN). O ROIAlign aborda limitações presentes no método tradicional de agrupamento de Regiões de Interesse (ROI) ao empregar interpolação bilinear durante o processo de agrupamento. Esse aprimoramento mitiga problemas de desalinhamento, garantindo a captura precisa de informações espaciais do mapa de recursos de entrada. Como resultado, há um aumento na precisão da segmentação, especialmente para objetos de pequenas dimensões.

A FPN desempenha um papel crucial na extração de recursos ao construir uma pirâmide de recursos em várias escalas, uma abordagem envolve a integração de características provenientes de diversas escalas, formando uma pirâmide que capacita o modelo a obter uma compreensão mais abrangente do contexto do objeto. Essa ampla gama de escalas facilita uma detecção e segmentação de objetos mais eficazes, abrangendo uma variedade significativa de tamanhos de objetos. Em essência, a FPN enriquece a representação de características, permitindo ao modelo lidar de maneira mais eficiente com objetos de diferentes dimensões. A Figura 13, mostra a estrutura Mask R-CNN para segmentação de instâncias.

FIGURA 13 - ESTRUTURA MASK R-CNN PARA SEGMENTAÇÃO DE INSTÂNCIAS



Fonte: Adaptado do He et al. (2017).

3.2.3 Avaliação do desempenho dos modelos de detecção e segmentação

Para avaliação do desempenho dos modelos foram determinadas métricas utilizadas para obter uma avaliação quantitativa de detecção e segmentação, destaca-se o mAP (precisão média) e a Interseção sobre União (IoU), descritas nas Equações (6 e 7), respectivamente. O mAP incorpora a métrica de Precisão (P) e Recall (R). Isso torna o mAP uma métrica adequada para avaliar a maioria das aplicações de detecção e segmentação de objetos.

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N \text{AP}_i \quad (6)$$

$$\text{IoU} = \frac{\text{Área de Sobreposição}}{\text{Área da União}} \quad (7)$$

onde AP_i é a precisão média das classes e N é o número de classes. A métrica de Precisão Média (AP) está vinculada à Interseção sobre União (IoU), conforme expresso na Equação 7, e ilustra como as caixas se sobrepõem na detecção de objetos. Essa métrica proporciona uma caixa de saída que envolve perfeitamente os objetos, possibilitando a avaliação dos erros de localização em modelos de detecção de objetos. Nesses modelos, a célula da grade é encarregada de prever as caixas delimitadoras e suas pontuações de confiança. Portanto, $\text{IoU} = 1$ quando a caixa delimitadora prevista é idêntica à caixa real.

Outra métrica empregada para obter uma avaliação quantitativa da segmentação é o cálculo da Média do Erro de Classificação (ME), conforme descrito pela Equação 8:

$$\text{ME} = 1 - \frac{|B_O \cap B_T| + |F_O \cap F_T|}{M \times N} \quad (8)$$

A imagem foi segmentada em primeiro plano e segundo plano. O primeiro plano é o alvo do experimento e o resto é o plano de fundo. B_O é o número de pixels do fundo da imagem da verdade fundamental; B_T é o número de pixels do fundo segmentado; F_O é o número de pixels do primeiro plano da imagem de verdade do solo; F_T é o número de pixels do primeiro plano segmentado pelo modelo e; $M \times N$

significa o total de pixels da imagem de teste. Assim, quanto menor for o ME, melhor será o resultado da segmentação.

3.2.4 Classificação

Após a detecção e segmentação das folhas de erva-mate foi proposta uma abordagem híbrida para a tarefa de multiclassificação automática de doenças de plantas usando imagens digitais RGB digitais, combinando os descritores de extração de características (Momento Hu, Texturas Haralick e Histograma de Cores) e rede neural convolucional.

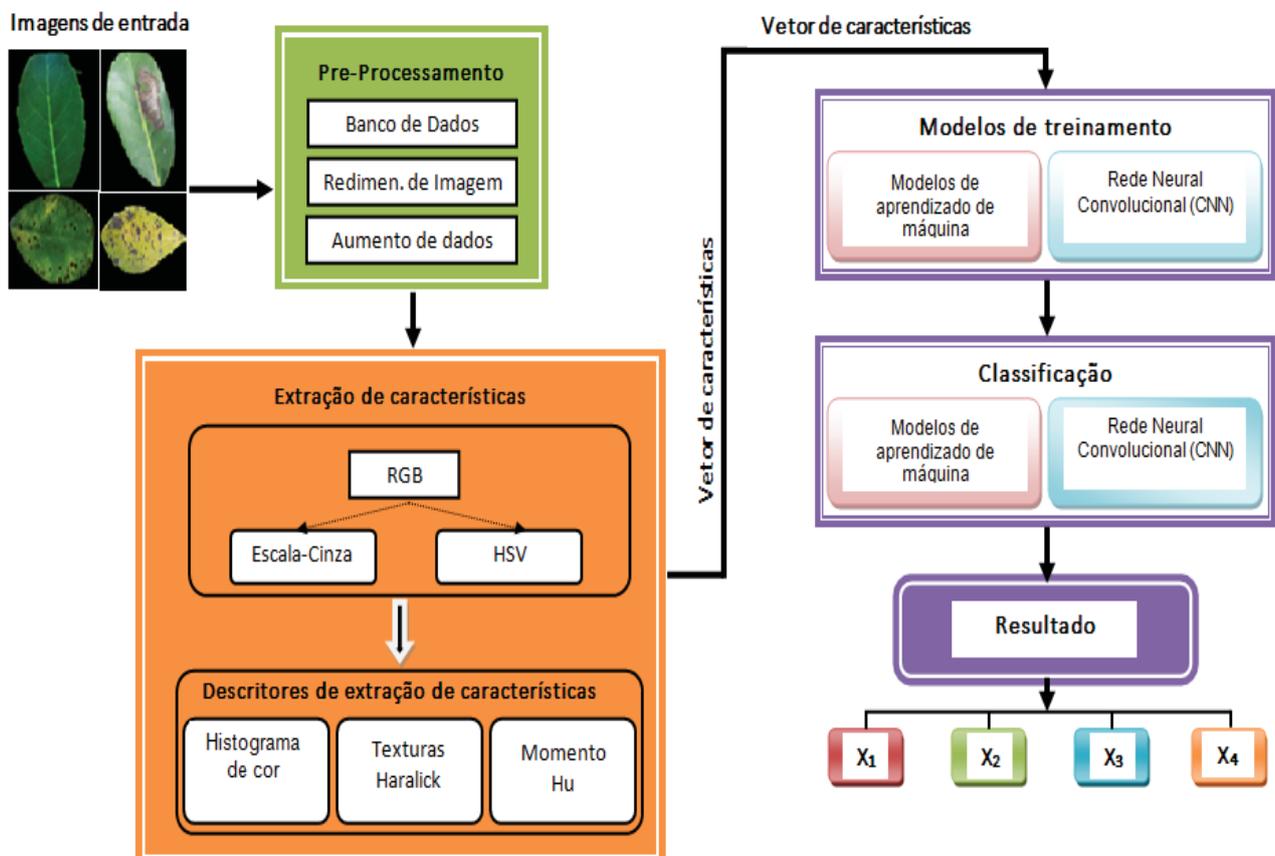
A abordagem proposta descrita no fluxograma metodológico na Figura 14, compreende 3 seções. Na primeira seção, as imagens digitais RGB foram submetidas a etapa de pré-processamento que inclui o redimensionamento das imagens a fim de facilitar a análise em uma única escala. Como os métodos baseados em aprendizado profundo necessitam de uma grande quantidade de imagens foi necessário aumentar as amostras disponíveis por meio de processamento de imagens.

Na segunda seção converte-se o espaço de cores da imagem de BGR (Blue-Green-Red) para as imagens RGB aprimoradas - essas são convertidas em tons de cinza para identificar a região infectada e a imagem de entrada foi convertida para o espaço de cores HSV (*Hue-Saturation-Value*) para que a área infectada pudesse ser facilmente segmentada em toda a imagem. Portanto, são criadas as máscaras para extrair áreas específicas da imagem com base em faixas de cores (verde e marrom). A clusterização é outra técnica importante aplicada para identificação de Região de Interesse (ROI). O agrupamento k-means foi empregado para obter ROI nas imagens de folhas de plantas. O k-means agrupa pixels semelhantes em clusters, segmentando assim a imagem em diferentes regiões com base em cores, intensidades ou outras características de pixel. Ao aplicar o k-means, é possível identificar agrupamentos naturais na imagem, contribuindo para a identificação eficiente de regiões relevantes. O agrupamento consegue segmentar imagens em

diferentes regiões com base em cores, intensidades ou outras características de pixel. Essas operações foram realizadas como parte do pré-processamento de imagens, para posterior extração de características.

Na terceira seção, aplicaram-se três técnicas de extração artesanal de características baseadas em Momento Hu, Texturas Haralick e Histograma de Cores. Esse conjunto reduzido de características, resume a maior parte da informação contida na imagem original. Finalmente, o vetor de características foi criado por meio da fusão dessas características, com o objetivo de implementar modelos de classificação. Esses modelos incluem a utilização de uma rede neural convolucional do tipo CNN, juntamente com três algoritmos de aprendizado de máquina: Random Forest (RF), Support Vector Machine (SVM) e K-Nearest Neighbors (KNN). Os resultados da classificação correspondem às categorias das folhas, especificamente classificando as patologias associadas a cada uma delas.

FIGURA 14 - FLUXOGRAMA METODOLÓGICO PROPOSTO



Fonte: Autor (2023).

Em geral, no presente trabalho, o processo de implementação é dividido em duas fases principais: a fase de treinamento e a fase de teste. Na fase de treinamento, as imagens são inicialmente pré-processadas, convertendo imagens RGB para para espaço de cor HSV. Posteriormente, são extraídas características relevantes das imagens e selecionada a melhor combinação de atributos para o processo de classificação. Durante a fase de teste, imagens não vistas anteriormente pelo modelo são selecionadas e inseridas no sistema. O modelo é consultado com essas imagens, que são então classificadas em relação à doença específica e à classe correspondente da folha de erva-mate.

3.2.5 Extração de características

Para quantificar os pontos de interesse na imagem, utilizou-se descritores globais e locais. A cor do sintoma na folha é analisada por meio de um histograma de cores, enquanto a forma é quantificada usando momentos Hu e a textura é avaliada através de um método Haralick. A extração dessas características é fundamental, pois a cor, forma e textura são propriedades essenciais que predominam em diversas patologias no conjunto de dados. O processo de extração de características é realizado individualmente para cada imagem, onde os três tipos de características são extraídos, combinados em um único vetor e armazenado no formato HDF5 para análise posterior. Assim, o modelo proposto no estudo baseou-se na combinação de três descritores artesanais para extração de características: Momento Hu, Textura Haralick e Histograma de Cores. Esses descritores foram intergados na arquitetura CNN.

A seguir, serão descritas as abordagens de extração de características implementadas na pesquisa:

a) *Histograma de Cores (descritor de cor)*: O histograma baseado na cor é uma técnica utilizada para determinar a distribuição de cores em uma imagem. A cor representa um fator de grande relevância na classificação de imagens, como é o caso

da identificação de doenças em plantas. O histograma de cores reflete a quantidade de valores de pixels presentes em compartimentos individuais, correspondentes às cores listadas no espectro cromático. As frequências acumuladas de ocorrência de cada cor compõem o vetor de características. Por exemplo, tons de vermelho, verde e azul são considerados em imagens coloridas RGB. O comprimento do vetor de características resultante será de $8 \times 8 \times 8 = 512$, caso se opte, por exemplo, por um histograma com 8 intervalos para cada canal de cor.

b) Textura Haralick (descriptor de textura): A textura de uma imagem é uma propriedade que descreve a superfície e a aparência do objeto na imagem. A textura pode ser categorizada em dois grupos: textura regular e textura estocástica, com base no nível de aleatoriedade. A textura regular refere-se ao arranjo ordenado dos componentes básicos de um objeto, enquanto a textura estocástica envolve a disposição aleatória desses componentes básicos. Devido às distribuições aleatórias de células no tecido da folha, necessita-se de cálculo de características estatísticas para realizar a análise de textura das imagens, elas apresentam em sua maioria uma textura estocástica. O GLCM é um meio de considerar a proporção espacial de pixels para calcular essas características estatísticas. Haralick et al. (1973) propuseram como o GLCM poderia ser usado na quantificação de textura. Isso tem relevância na identificação de doenças em plantas, uma vez que a análise da textura pode oferecer *insights* valiosos sobre as alterações patológicas. No descriptor Textura Haralick, a matriz de coocorrência de nível de cinza é calculada usando pixels em tons de cinza i e j , que são expressos como o número de matrizes de coocorrência em diferentes direções, conforme representado abaixo, (Equação 9):

$$p(i, j|d, \theta) = \frac{p(i, j|d, \theta)}{\sum_i \sum_j p(i, j|d, \theta)} \quad (9)$$

onde $p(i, j)$ é a matriz de frequências relativas, d é a distância de pixel que ocorre dentro de $(x1, y1)$ e θ é a direção do pixel $(x1, y1)$. A matriz $p(i, j)$ fornece os valores de probabilidade estatística para as transições entre os níveis de cinza i e j em uma determinada distância d e ângulo θ . Essa expressão é fundamental para modelar as relações estatísticas entre os pixels em termos de intensidade, distância e orientação. Portanto, características texturais são extraídas da dependência espacial do tom

cinza, como: Contraste, Correlação, Entropia e Energia de Homogeneidade. Detalhe de cada característica serão descritas a seguir:

Contraste - mede a frequência espacial de uma imagem e é um momento diferente do GLCM (Equação 6). É a diferença entre o valor mais alto e o mais baixo do conjunto adjacente de pixels. A textura de contraste mede as variações locais presentes na imagem. Uma imagem com baixo contraste apresenta termo de concentração GLCM em torno da diagonal principal e apresenta baixas frequências espaciais, (Equação 10):

$$\text{Contraste} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - j)^2 \quad (10)$$

Correlação - avalia a relação linear entre o tom de cinza de um pixel e seus vizinhos em toda a imagem, analisa as dependências lineares presentes na distribuição dos tons de cinza, (Equação 11):

$$\text{Correlação} = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j) p(i,j)}{\delta_i \delta_j} \quad (11)$$

onde δ_i, δ_j e μ_i, μ_j são os desvios padrão e médias de p_i e p_j , respectivamente.

Entropia - é uma métrica que avalia a complexidade da imagem, sendo sensível a padrões complexos e variações nos níveis de cinza. À medida que a complexidade da imagem aumenta, a entropia tende a aumentar, (Equação 12):

$$\text{Entropia} = \sum_i \sum_j p(i, j) \quad (12)$$

Energia - é definida como a soma dos quadrados dos elementos da GLCM, permanecendo inalterada por padrão, (Equação 13):

$$\text{Energia} = \sum_{i,j} (i, j)^2 \quad (13)$$

c) *Momentos Hu (descritor de forma)*: Os momentos invariantes de Hu são descritores de forma global, uma abordagem na visão computacional proposto por Ming-Kuei em 1962. Normalmente, são utilizados para obter os atributos relacionados ao formato das manchas visíveis nas imagens. O momento Hu é um método que pode determinar

formas ou padrões independentemente da escala, localização e orientação, ou seja, esses momentos invariantes de forma são utilizados para descrever características globais de objetos em uma imagem, oferecendo uma representação distintiva que permanece constante independentemente de rotações, translações ou escalas. Os momentos geométricos podem ser descritos no processamento de imagens como um momento de uma função $f(y, z)$ de ordem $(r + s)$ e $(r, s = 0, 1, 2, \dots, \infty)$. Hu especifica sete desses momentos, que são obtidos usando o momento central da sequência três e são chamados coletivamente de momento Hu. Quando o momento Hu é aplicado a uma imagem, um vetor de características de tamanho 7-d é encontrado. Podem ser descritos no processamento de imagens da seguinte (Equação 14):

$$M_{rs} = \int_{p_1}^{p_2} \int_{q_1}^{q_2} y^r z^s f(y, z) dy dz \quad (14)$$

onde, M_{rs} - momento de uma função $f(y, z)$ de ordem $(r + s)$ e $(r, s = 0, 1, 2, \dots, \infty)$ e y, r, z, s - definem a verdadeira função de momento básico. Os momentos invariantes são uma combinação não linear de momentos rotativos invariantes.

A extração de características usando descritores visa encontrar um subconjunto de características mais relevantes para modelos de classificação final.

a) *Normalização*: A normalização é crucial ao empregar algoritmos de aprendizado de máquina que não fazem suposições específicas sobre a distribuição dos dados. No presente estudo, utilizou-se a função *StandardScaler* da biblioteca *scikit-learn*, uma técnica comum de transformação de dados no campo de aprendizado de máquina para padronizar a escala de características ou variáveis numéricas. Essa abordagem utiliza a média e o desvio padrão para redimensionar os dados, resultando em características que seguem uma distribuição com média zero e variância unitária, o que facilita a comparação e interpretação dos dados (Singh & Singh, 2020). Cada instância, $x_{i,n}$ nos dados é transformada em $z_{i,n}$ (o valor normalizado), contribuindo para uma representação consistente e comparável no contexto do estudo, (Equação 15):

$$z_{i,n} = \frac{(x_{i,n} - \mu_i)}{\sigma_i} \quad (15)$$

onde μ_i e σ_i denotam a média e o desvio padrão dos valores dos recursos no conjunto de dados, respectivamente.

3.2.6 Aprendizado Profundo

d) *Rede Neural Convolutacional*: as redes neurais profundas, incluindo as do tipo CNN, são os modelos mais utilizados para processamento de imagens, principalmente para identificação e classificação de doenças em plantas (Golhani et al., 2018; Zhang et al., 2019; Abdu et al., 2020; Jaiganesh et al., 2020; Qi et al., 2021).

Neste trabalho, técnicas de extração de características (Momento Hu, Texturas Haralick e Histograma de Cores) e Rede Neural Convolutacional (CNN) são empregadas para multiclassificação de patologias em plantas, usando um conjunto de dados específico composto por folhas de erva-mate com sintomas visíveis nas imagens ópticas. A saída é a confirmação da doença na folha e a probabilidade de identificação correta. Formalmente, a filtragem realizada em cada camada neural convolutacional pode ser escrita de acordo com as Equações 16 e 17.

$$s = \sum_{i=1}^{nv} x_i p_i + b \quad (16)$$

$$y = f(s) \quad (17)$$

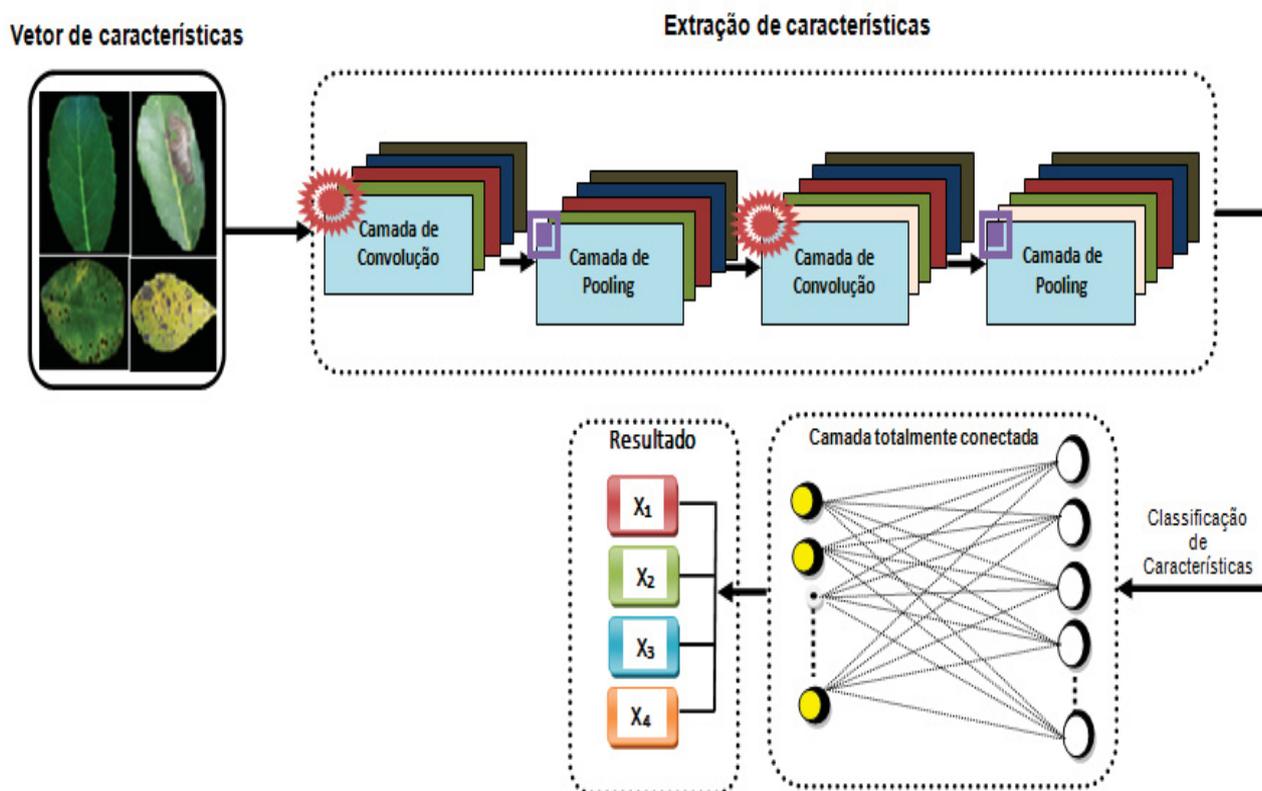
onde x_i representa o valor de um pixel na vizinhança, p_i seu respectivo peso e b uma constante adicionada (bias).

A classificação bem sucedida de características significativas da imagem é altamente dependente da determinação adequada dos pesos mais apropriados para a tarefa desejada. Esses pesos são estimados com base em amostras de treinamento de forma supervisionada. Os pesos da rede foram otimizados por um algoritmo de otimização denominado Adaptive Moment Estimation (*Adam*) e realizado o *dropout*, método de regularização para redes neurais que ajuda a diminuir o aprendizado interdependente entre os neurônios e o *dropout* para reduzir o ajuste excessivo, adicionando uma penalidade à função de perda.

A arquitetura da rede proposta foi projetada para aprender funcionalidades usando conexões densas de uma rede neural feed-forward (conhecida como rede neural de alimentação) e realizar classificações por meio da função *softmax*. Cada fase de treinamento e teste dos dados foi realizada separadamente, para evitar problemas de superajuste foi aplicado aumento nos dados de treinamento. Um *dropout* de 0,2% foi implementado para mitigar o *overfitting* do modelo. O algoritmo otimizador *Adam* foi empregado. As funções de ativação adotadas foram, a função ReLU (*rectified linear unit*) e a função *softmax*, utilizando uma taxa de aprendizagem padrão e um fator de decaimento.

Em geral, na arquitetura, a camada de entrada foi equipada com 96 neurônios, utilizando *ReLU* como função de ativação e uma taxa de desistência (*dropout*) de 20% (0,2). As três camadas densas subsequentes possuíam 128, 64 e 32 neurônios, respectivamente, cada uma com uma taxa de *dropout* de 20% e *ReLU* como função de ativação. Finalmente, a última camada, que é a camada de saída, consiste em 8 neurônios, usando *Softmax* como função de ativação e *Adam* como otimizador. O diagrama de estrutura do classificador DNN proposto é ilustrado na Figura 15.

FIGURA 15 - VISÃO GERAL DA ARQUITETURA CNN



Fonte: Autor (2023).

3.2.7 Comparação com outros classificadores

Para fins de comparação, o mesmo conjunto de dados foi utilizado como entrada para outros três classificadores convencionais, descritos a seguir:

a) *Random Forest (RF)*: método de aprendizagem supervisionada proposto por Ho em 1995, que pode ser utilizado para resolver problemas de classificação, regressão e outros. O algoritmo consiste em uma combinação de classificadores de árvores de decisão que são comumente treinados para atingir o objetivo proposto. As árvores podem produzir resultados diferentes com base nas diferenças de entrada. O resultado desta etapa é a combinação dos diferentes resultados. Cada árvore dá um voto unitário para a classe mais popular para classificar um vetor de entrada (Qi et al.,

2021). Quanto maior o número de árvores na floresta, maior será a precisão e menor será a chance de *overfitting*. O algoritmo possui duas etapas: a primeira gera a árvore aleatória e a segunda produz uma predição usando o classificador de árvore aleatória que foi criado na primeira etapa (Khilari et al., 2021). A combinação de árvores é feita usando *bootstrap bagging* junto com *feature random bagging*. As amostras de treinamento são selecionadas aleatoriamente usando a técnica de *bootstrap bagging*, $D = \{d_1, d_1, d_1, \dots, d_n\}$, com valores $Y = \{y_1, y_2, y_3, \dots, y_n\}$, o processo de empacotamento é feito repetidamente para identificar rótulos de classe para registros de dados com base em valores de atributos. O processo de treinamento é repetido até que o resultado ideal seja obtido (Shingade et al., 2022). Assim, as previsões globais resultantes do classificador RF podem ser dadas pela, (Equação 18):

$$\hat{f} = \frac{1}{\beta} \sum_{b=1}^{\beta} f_b(d') \quad (18)$$

onde, $b = \{1, 2, \dots, \beta\}$ - indica o número de ensacamento realizado, β - é o número máximo de ensacamento, d' - indica a amostra não vista, e $f_b(d')$ - indica a previsão feita na amostra. Como muitas árvores são construídas em RF, todas as árvores são finalmente agregadas. A influência do ruído e de outras distorções numa determinada árvore pode degradar o desempenho de outras árvores. O desvio padrão de previsão ainda é calculado para todas as árvores individuais para obter resultados precisos (Equação 19), Shingade et al. (2022):

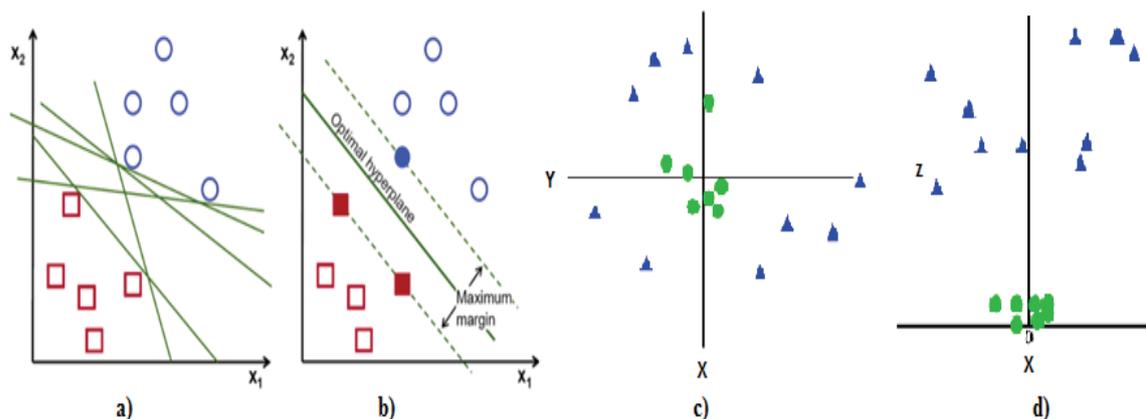
$$\sigma = \sqrt{\frac{\sum_{b=1}^{\beta} (f_b(d') - \hat{f})^2}{\beta - 1}} \quad (19)$$

Na presente pesquisa, as características das folhas foram utilizadas como variáveis de entrada para este algoritmo. Cada árvore do modelo foi treinada com um subconjunto aleatório dessas características. Durante o processo de treinamento, o algoritmo cria divisões entre as características que minimizam o erro de classificação (Buttan et al., 2021). Cada árvore é construída diferente das outras árvores. Após o treinamento, as folhas foram classificadas de acordo com suas características de insumo. Assim, cada árvore produziu uma predição, e a classe final foi determinada pela maioria das predições.

b) Support Vector Machine (SVM): técnica de aprendizado de máquina supervisionado proposta por Vapnik e Chervonenkis em 1963, baseada em Kernel com alto poder computacional para resolver problemas de classificação e regressão, normalmente utilizada para dados que são separáveis linearmente ou não linearmente. O modelo aplicado no estudo é não linear, pois o conjunto de dados possui diversas funcionalidades. O método *Kernel trick* foi usado para encontrar o hiperplano que melhor separa as classes em um espaço de alta dimensão usando planos ou funções matemáticas. Portanto, um hiperplano pode ter múltiplas linhas ou limites de decisão, e os vetores de suporte são os pontos que estão mais próximos do hiperplano (Khilari et al., 2021).

Este modelo, no entanto, é usado principalmente para resolver dificuldades de categorização em um espaço dimensional superior ou infinito (Koklu e Ozkan, 2020), e pode converter um problema inseparável em um separável. O objetivo deste algoritmo é encontrar o limite de decisão ideal para dividir o espaço multidimensional de variáveis de acordo com classes, de modo que novas amostras possam ser prontamente classificadas na categoria correta. Para melhor compreensão, apresentou-se as figuras 16a e 16b que representam um conjunto de dados separável em um espaço 2-d (x e y) ao aplicar uma linha reta. A Fig. 16c ilustra uma situação não linearmente separável; Para separar essas classes, uma terceira dimensão (z) pode ser adicionada (conforme exibido na Figura 16c), por exemplo ($z = x^2 + y^2$). Os diferentes espaços de características do modelo SVM são do tipo linear (exemplo a e b) e do tipo não linear (exemplo c e d). Os pontos representados nas figuras descrevem diversas classes do conjunto de dados, distribuídas ou separados em espaços distintos.

FIGURA 16 - DIFERENTES ESPAÇOS DE CARACTERÍSTICAS DO MODELO SVM



Fonte: Adaptado do Khilari et al. (2022).

c) *K-Nearest Neighbours (KNN)*: método desenvolvido inicialmente por Evelyn Fix e Joseph Hodges em 1951, caracterizado como um classificador não paramétrico convencional, que tem sido utilizado como classificador em muitos problemas de regressão e classificação de padrões. KNN funciona encontrando os k pontos de dados mais próximos de um novo ponto de dados e, em seguida, usando a classe desses pontos para prever a classe do novo ponto. No caso deste estudo, o algoritmo permitiu a categorização de dados que prevê a probabilidade de uma característica dos dados pertencer a uma determinada classe de folha. A distância entre pontos de dados ou feições pode ser medida usando diferentes métricas, como distância euclidiana ou distância de Manhattan (Equação 20), onde com base nos dados de aprendizagem existentes, novos dados de entrada são atribuídos ao cluster mais próximo da amostra estabelecida (Chen et al., 2018; Koklu e Ozkan, 2020; Khatri et al., 2022). Suponha dois fatores $X = (x_1, x_2, x_3, \dots, x_n)$ e $Y = (y_1, y_2, y_3, \dots, y_n)$:

$$d(X, Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (20)$$

Na tarefa de classificação também utilizou-se a validação cruzada (*cross-validation, em inglês*). É uma técnica estatística comumente utilizada em aprendizado de máquina para estimar o desempenho de um modelo. Para a escolha de melhor desempenho do modelo foram realizado 5 tipos de validação cruzada. Portanto, em cada rodada de treinamento, usando os 80% dos dados separados anteriormente, os dados foram particionados usando-se 10, 20, 50, 100-*fold* cross-validation. No caso

do primeiro teste, os dados são separados em 10 grupos distintos e, a cada rodada de treinamento, um desses grupos é usado como teste e os outros nove usados como treino.

3.2.8 Avaliação do desempenho dos modelos de classificação

Para avaliar o desempenho dos modelos de classificação, foram calculadas as seguintes métricas: precisão, recall, medição do F-score e exatidão (Equações 21, 22, 23 e 24). Essas métricas foram calculadas a partir do número de resultados verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN).

As medidas de precisão e exatidão verificam a capacidade do classificador de replicar os mesmos resultados precisos em iterações contínuas (Ma et al., 2018; Turkoglu e Hanbay, 2019; Abdu et al., 2020). A *Recall* avalia a capacidade do método de detectar com sucesso classes classificadas como positivas para todas as observações (a precisão do produtor). *F-score* é a média harmônica de recall e precisão (Tharwat, 2020), comumente usada para avaliar sistemas de classificação binária, que classificam os exemplos como “positivos” ou “negativos”. As métricas de precisão e pontuação F são indicadores abrangentes, portanto, um valor maior significa maior precisão, enquanto recall e precisão estão relacionados a erros de omissão e comissão, respectivamente.

$$\text{Precisão} = \frac{VP}{VP+FP} \quad (21)$$

$$\text{Recall} = \frac{VP}{VP+FN} \quad (22)$$

$$\text{F - score} = \frac{2VP}{2VP+FP+FN} \quad (23)$$

$$\text{Acurácia} = \frac{VP+VN}{VP+FP+FN+VN} \quad (24)$$

Outra forma considerada para avaliar o desempenho dos modelos é com base na curva ROC (Receiver Operating Characteristic) e sua respectiva área sob a curva,

conhecida como AUC (Area Under the Curve). A construção da curva ROC e AUC, foi uma parte crucial da tarefa de classificação (ver Figura 23). A curva ROC é uma representação gráfica do desempenho de um classificador em diferentes limites de discriminação. Ela é criada plotando a sensibilidade (taxa de verdadeiros positivos) contra a especificidade (1 - taxa de falsos positivos) para vários valores de limite de decisão. Em outras palavras, a sensibilidade representa a proporção de casos positivos reais que são corretamente identificados pelo classificador, enquanto a especificidade representa a proporção de casos negativos reais que são incorretamente classificados como positivos. A área sob a curva (AUC) é um indicador numérico do desempenho do classificador: quanto maior a AUC, melhor o modelo está em distinguir entre as classes positivas e negativas.

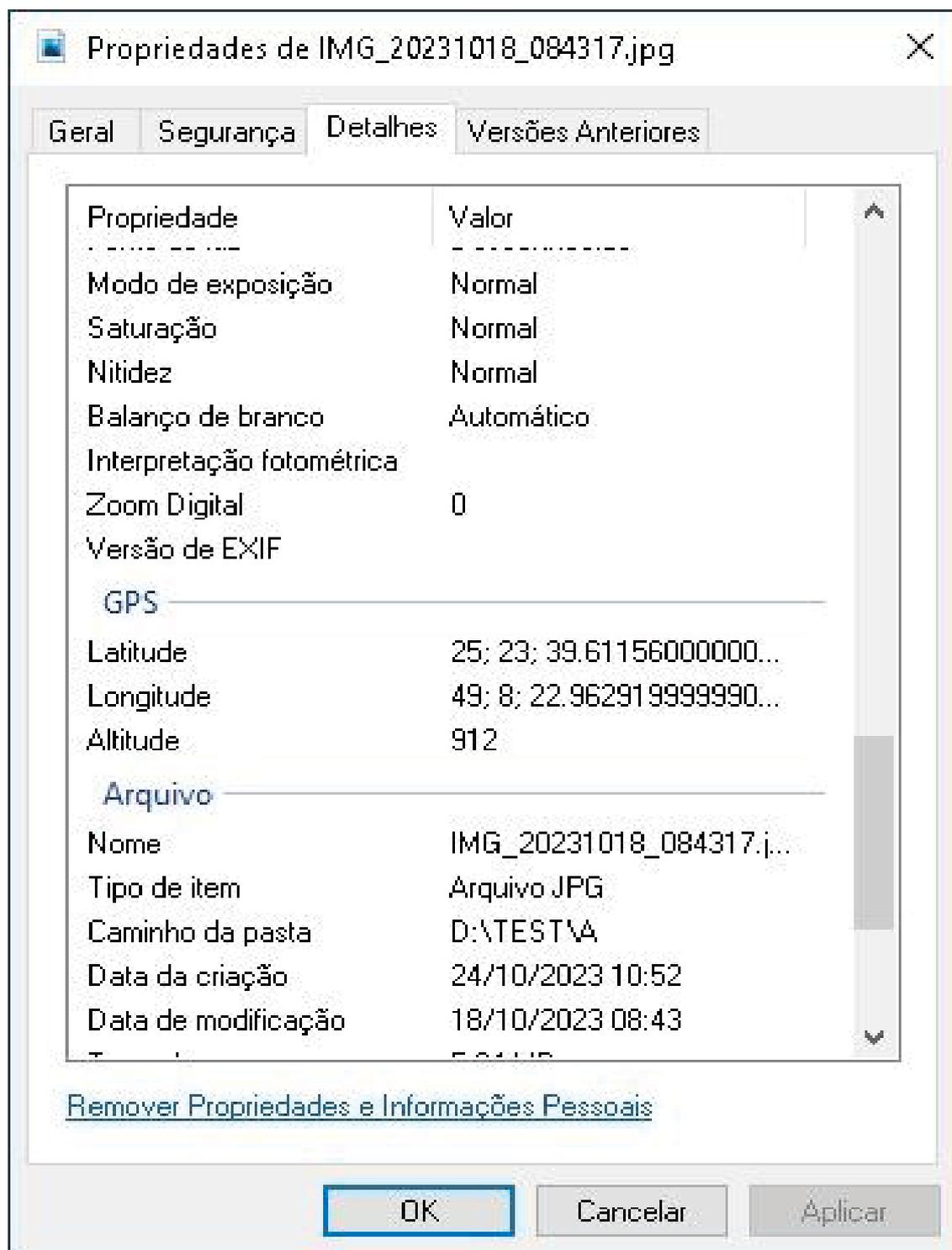
A tarefa de classificação das doenças em folhas de erva-mate foi implementada em linguagem Python, dentro do ambiente Google Colab, usando a linguagem Python e explorando os recursos de Scikit-learn e TensorFlow.

3.3 MAPEAMENTO DA OCORRÊNCIA DE PATOLOGIAS

Com base nas fotografias digitais das folhas de erva-mate presentes no conjunto de dados de treinamento, mapeamos a ocorrência de patologias utilizando os metadados *Exchangeable Image File Format* (EXIF) incorporados nas imagens, conforme ilustrado na Figura 17.

As imagens com coordenadas GPS (Global Positioning System) foram transferidas para um Sistema de Informação Geográfica (SIG) para possibilitar seu mapeamento. Nesse contexto, foi utilizado o *plugin Geotag e Import Photos* do Quantum GIS - pode ser acessado em: <https://plugins.qgis.org/plugins/ImportPhotos/>. Essa ferramenta permite a importação de fotos com marcação geográfica (formatos jpg ou jpeg) como pontos para o QGIS. O usuário pode selecionar uma pasta contendo fotos, e apenas as imagens com marcação geográfica serão incluídas. Posteriormente, uma camada foi criada, incluindo informações como nome da imagem, diretório, data e hora da captura, altitude, longitude, latitude, azimute, direção norte, fabricante e modelo da câmera, título, comentário do usuário e caminho relativo. A ferramenta possui dois botões: um para importar fotos com georreferenciação e outro que permite clicar em um ponto no mapa para exibir a foto correspondente, juntamente com informações de data, hora e altitude. Além disso, o usuário pode criar diferentes tipos de arquivos, como GeoJSON, SHP, GPKG, CSV, KML e TAB.

FIGURA 17 -EXEMPLO DE METADADOS DE UMA IMAGEM



Fonte: Autor (2023).

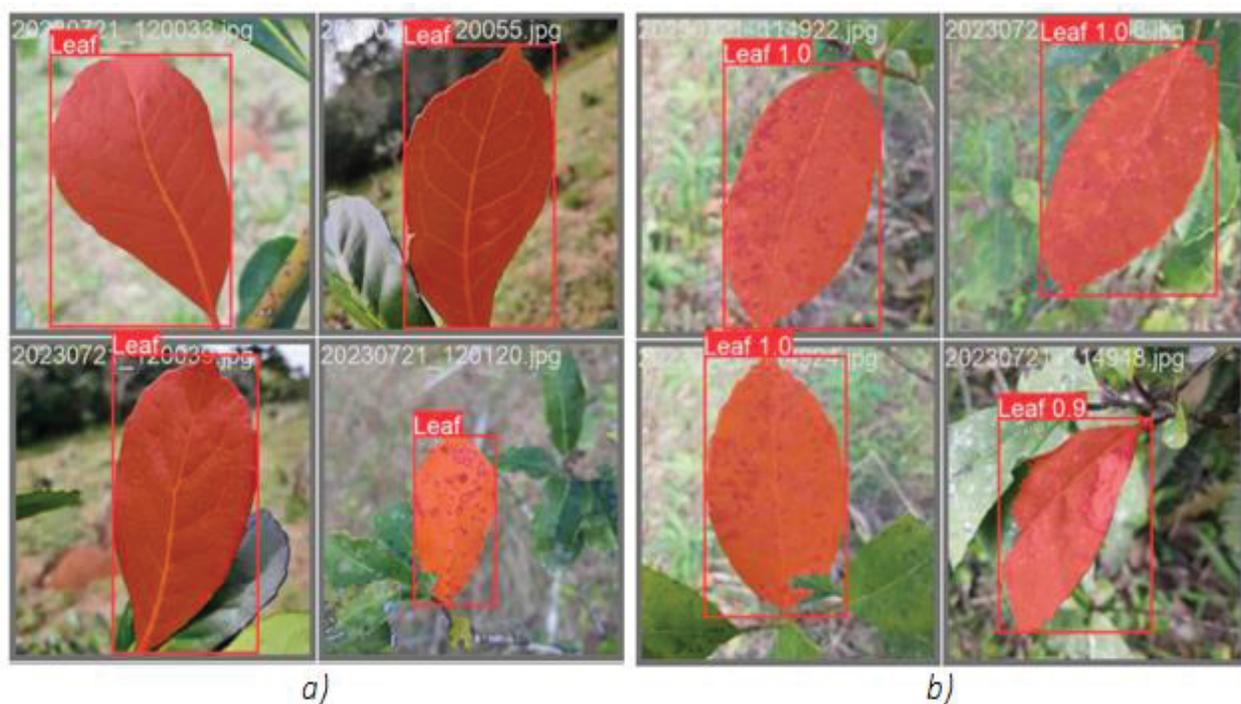
4 RESULTADOS

Esta seção apresenta os resultados obtido nos experimentos de detecção, segmentação e classificação:

4.1 DETECÇÃO DA FOLHA DE ERVA-MATE

Durante o treinamento do modelo, identificou-se como alvos de detecção e segmentação as folhas relativamente maiores presentes na imagem, enquanto as folhas relativamente menores no fundo das imagens não foram consideradas alvos de detecção. Na Figura 18 é apresentado o resultado de detecção e segmentação de instâncias (a), onde é possível observar cada folha enquadrada. Em seguida é fornecida a probabilidade correspondente de reconhecimento correto da folha (b), expressa em forma de porcentagem.

FIGURA 18 - RESULTADO DE DETECÇÃO E SEGMENTAÇÃO DE INSTÂNCIAS

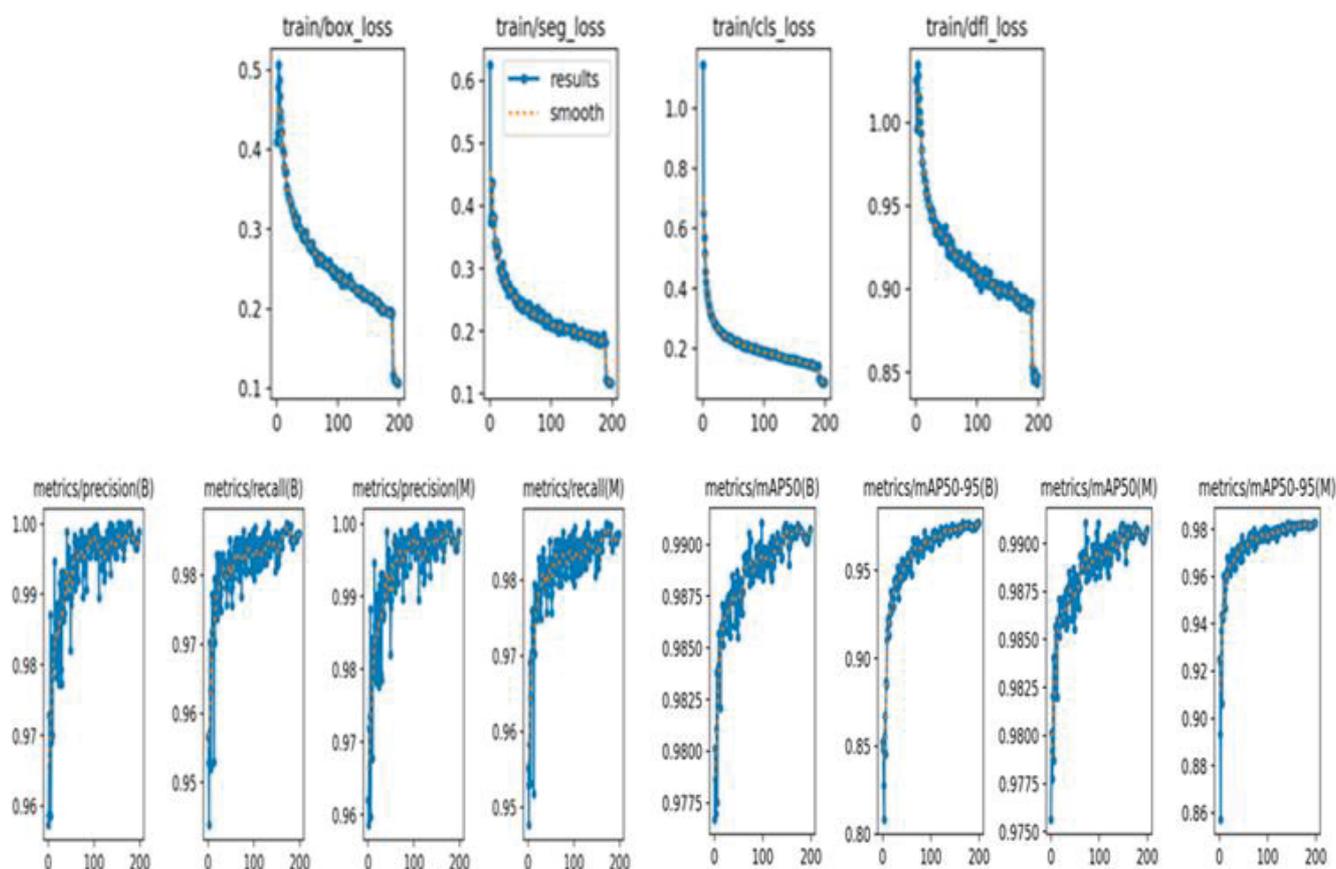


Fonte: Autor (2023).

Na Figura 19 é apresentada a evolução dos descritores de qualidade, após 200 épocas de treinamento. Durante o treinamento foram determinadas as curvas de valor da função de perda para os conjuntos de treinamento e validação, levando em consideração a perda do frame de detecção do objeto e a perda de classificação. De acordo com a disposição da esquerda para a direita, temos: a perda de caixa de regressão (*train/box_loss*), a perda de objeto (*train/obj_loss*), a perda de classificação (*train/cls_loss*) e as métricas de desempenho, precisão e *recall* para objetos de tamanho reduzido e médio. Também foram determinadas as precisões médias das detecções para objetos pequenos e médios considerando a sobreposição de 50% a 95% (IoU de 0,5 e 0,95). A sobreposição de 50% a 95% é uma faixa comum de valores usada para avaliar a precisão das detecções em problemas de detecção de objetos, especialmente em tarefas de detecção de múltiplos objetos em uma imagem.

A perda de caixa de detecção é um indicativo de quão bem o algoritmo consegue prever corretamente a classe de um objeto específico ou cobrir o objeto detectado com a caixa delimitadora prevista. Quanto menor a perda de caixa de detecção, maior a precisão na localização do objeto. A perda de objeto é uma medida da probabilidade de um objeto existir em uma região de interesse proposta. Um valor alto indica uma maior probabilidade de a janela da imagem conter um objeto. Essa medida auxilia na determinação da relevância da região em relação à presença do objeto. Quanto menor a perda de objeto, maior a confiança de que o objeto está presente na região. Uma perda de detecção menor reflete uma classificação mais precisa, ou seja, uma melhor capacidade do algoritmo de atribuir a classe correta ao objeto.

FIGURA 19 - VISUALIZAÇÃO DE TENDÊNCIA DE PERDA E DO DESEMPENHO DO MODELO PARA DETECÇÃO DA FOLHA



Fonte: Autor (2023).

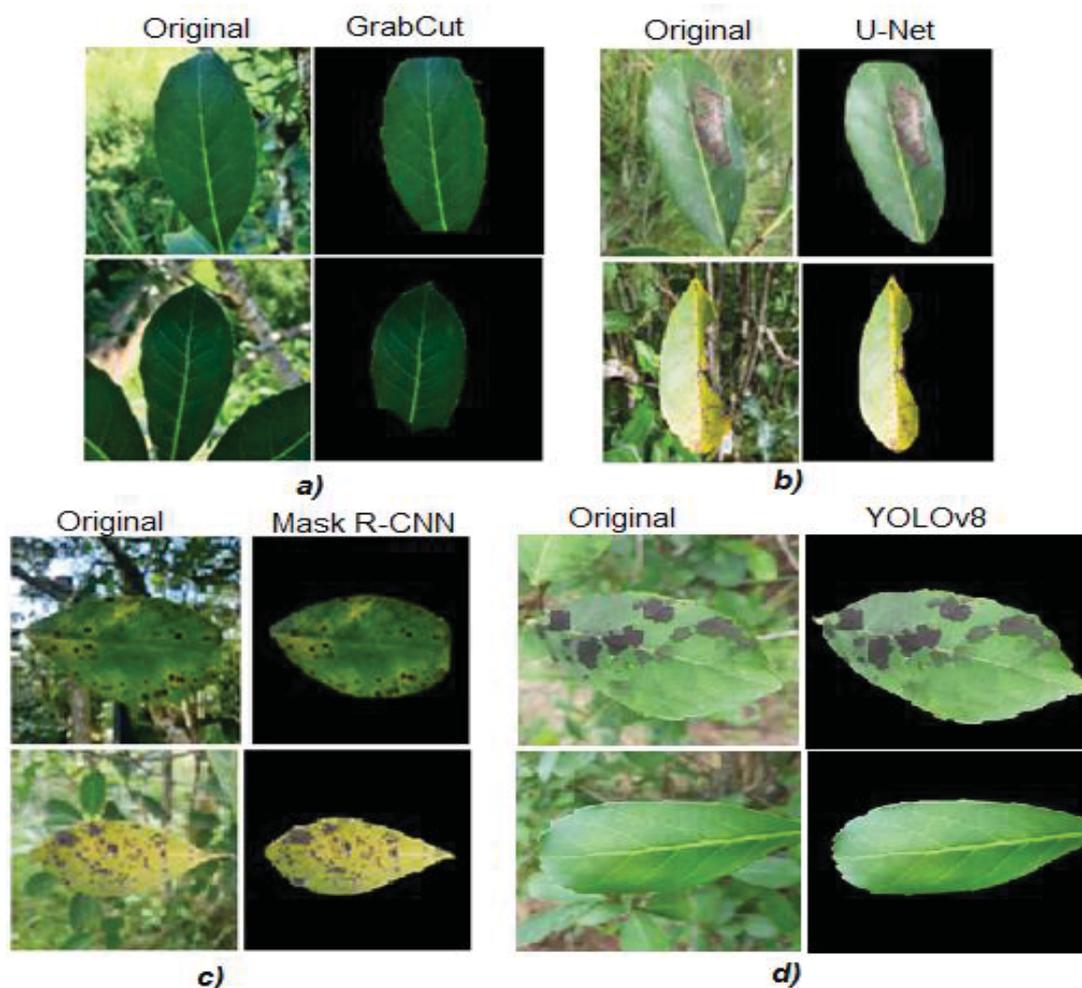
Durante o experimento, foi observada uma tendência de queda no valor da função de perda ao longo do processo de treinamento. Essa queda indica que a função de otimização (*Stochastic Gradient Descent-SGD*), foi capaz de otimizar a rede neural, atualizando constantemente os pesos e parâmetros da rede para melhor ajuste aos dados. O modelo utilizado no experimento foi composto por 195 camadas, o que indica uma arquitetura complexa capaz de processar informações em várias etapas. Além disso, foram utilizados 3.258.259 parâmetros, os quais foram ajustados durante o processo de treinamento.

O comportamento da redução gradual da função de perda durante o treinamento é indicativo de que a rede neural aproximou-se a um estado otimizado. Isso indica que o modelo está aprendendo e se tornando mais capaz de reconhecer e classificar corretamente os objetos (folhas).

4.2. SEGMENTAÇÃO DO FUNDO COMPLEXO DA FOLHA DE ERVA-MATE

Na Figura 20, são apresentados os resultados da segmentação do fundo complexo do conjunto de dados. O algoritmo YOLOv8 foi projetado com o propósito específico de identificar e isolar apenas as folhas-alvo do fundo. Para fins de comparação, também foram testados outros algoritmos, como o GrabCut, U-Net e Mask RCNN. Assim, por cada método testado (Figura 18), apresentou-se as imagens originais e a respectiva imagem segmentada, onde *a)* segmentação da folha com algoritmo tradicional GrabCut; *b)* segmentação com algoritmo U-Net; *c)* segmentação com algoritmo Mask RCNN e; *d)* segmentação com algoritmo YOLOv8.

FIGURA 20 - RESULTADOS DA SEGMENTAÇÃO DO FUNDO COMPLEXO DAS FOLHAS



Fonte: Autor (2023).

4.2.1 Comparação de diferentes métodos de segmentação

A Tabela 5 apresenta os resultados experimentais comparativos obtidos na tarefa de detecção e segmentação por meio da utilização dos diferentes métodos.

TABELA 5 - AVALIAÇÃO DE SEGMENTAÇÃO DAS FOLHAS DE ERVA-MATE

Método	Acurácia (%)	mAP50 (%)	ME (%)
GrabCut	89,29	-	24,65
U-Net	95,30	-	8,13
Mask R-CNN	98,9	97,4	4,28
YOLOv8n-seg	99,9	99,1	1,28

Fonte: Autor (2023).

O modelo YOLOv8n-seg demonstrou um desempenho satisfatório na tarefa de detecção e segmentação de objetos. Com uma precisão das caixas (bounding boxes) detectadas de 99,9% e recall de 98,6%, o modelo exibe alta acurácia na localização precisa dos objetos. Além disso, alcançou uma média de precisão (mAP50) de 99,1% com uma sobreposição de 50%, e uma mAP50-95 de 97,7% ao considerar uma variação de sobreposição de 50% a 95%. As máscaras de segmentação também obtiveram melhor desempenho, com precisão e recall 99,9% e 98,6%, respectivamente. Esses resultados indicam que o modelo tem a capacidade de detectar e segmentar das folhas de erva-mate com fundo complexo de forma precisa e em variedades escalas. Ainda o modelo apresentou um menor tempo de processamento, com um tempo médio de 40,3ms por conjunto de instâncias. Essa característica mostra a capacidade do modelo em realizar a segmentação e a redução dos requisitos computacionais necessários para o treinamento. O tempo médio dos outros modelos não foi estimado. Outro método que demonstrou desempenho satisfatório na segmentação foi o algoritmo Mask R-CNN, com um índice mAP de aproximadamente 97,4%. Isso coloca em foco especial a precisão das máscaras criadas para segmentação precisa das folhas.

A Tabela 5 também apresenta a medida de erro médio (EM), média para cada método. O algoritmo YOLOv8 revelou menor erro médio, seguido de Mask-RCNN, enquanto o algoritmo tradicional GrabCut apresentou o pior desempenho. Esses resultados indicam que quanto menor for o valor do EM, melhor será o resultado da segmentação.

4.3 CLASSIFICAÇÃO DE PATOLOGIAS USANDO DESCRITORES DE EXTRAÇÃO DE CARACTERÍSTICAS E CNN

Um dos objetivos do estudo é propor uma abordagem eficaz para detecção e diagnóstico de doenças com base nas imagens de folhas de erva-mate. Diversas tendências verificadas na literatura em termos de acurácia e precisões são discutidas, sendo algumas delas detalhadas na Tabela 5, por exemplo, a extração de características estruturais empregando diferentes técnicas para detectar patologias em plantas. Algoritmos de aprendizagem de máquina, incluindo RF, SVM, KNN e aprendizagem profunda, como CNN's, são também comumente empregados para classificar diferentes tipos de doenças.

4.3.1 Resultados de pré-processamento e segmentação

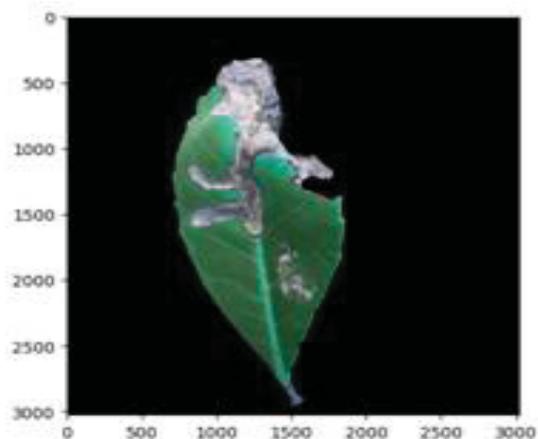
A Figura 21 ilustra as operações conduzidas durante a etapa de pré-processamento de imagem, fases consideradas essenciais para preparar a imagem original com objetivo de extração de características. Estas operações possibilitam a fácil segmentação das áreas infectadas em toda a imagem. As imagens originais do conjunto de dados, representadas em *a)*, passam por uma etapa essencial de aprimoramento, conforme mostrado em *b)*, onde são convertidas para o formato RGB. Esta fase é crucial para a análise de imagem, visando aprimorar a qualidade e a representação visual, facilitando a identificação e análise de características

específicas. Essa melhoria contribui diretamente para a precisão e eficiência dos algoritmos de extração de características, desempenhando um papel fundamental em tarefas como reconhecimento de padrões, detecção de objetos e análise de imagem em geral.

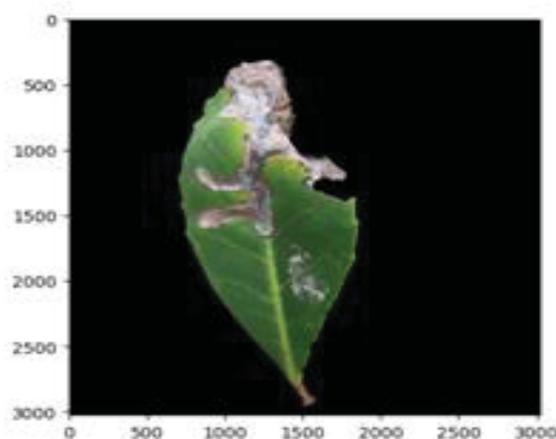
Em seguida, uma função é definida para receber a imagem aprimorada em RGB e convertê-la para o espaço de cores HSV (Hue, Saturation, Value), como ilustrado em c). A segmentação da imagem com base em cores é realizada por meio de uma função que utiliza informações de ambas as representações (RGB e HSV). Duas máscaras são criadas, uma para áreas verdes (indicativas de partes saudáveis da planta) e outra para áreas marrons (indicando áreas afetadas por doenças). Estas máscaras são combinadas para formar uma máscara final, destacando tanto as áreas saudáveis quanto as afetadas.

O resultado dessa segmentação é então aplicado à imagem original, conforme representado em d). Dessa forma, as diversas etapas do processamento, incluindo a conversão para o espaço de cores HSV e a segmentação com máscaras, resultam em uma imagem final que destaca as áreas saudáveis e afetadas da planta. A capacidade de segmentar as áreas afetadas na planta é evidenciada na imagem final, ilustrada em e).

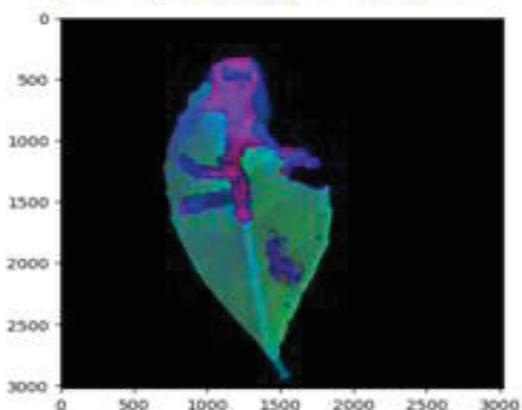
FIGURA 21 - RESULTADO DE PRÉ-PROCESSAMENTO E SEGMENTAÇÃO DE MANCHAS VISÍVEIS DE DOENÇAS



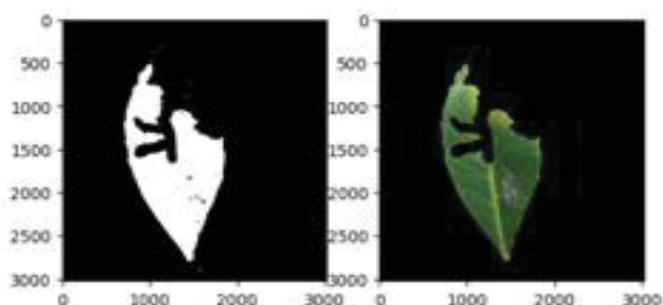
a) *Imagem original de folha*



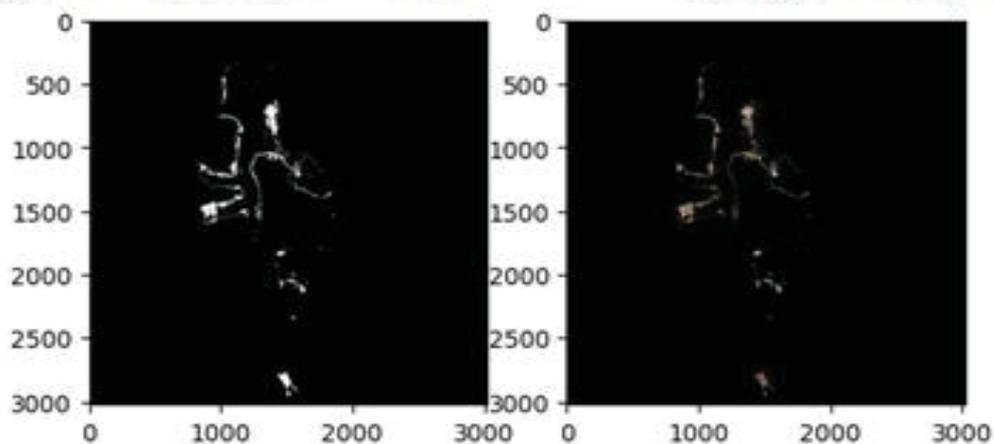
b) *Imagem RGB aprimorada*



c) *Imagem convertida para cor HSV*



d) *Imagem segmentada*



e) *Áreas infectadas*

Fonte: Autor (2023).

4.3.2 Delineamento de experimentos com base em descritores

A Tabela 6 apresenta os resultados do experimento de extração de características, destacando as dimensões do vetor que são representadas pelo número de exemplos ou instâncias do conjunto de dados, número de características ou atributos em cada exemplo, e o tempo necessário para cada conjunto de descritores. Cada experimento emprega um conjunto específico de descritores para extrair características das 3728 instâncias no conjunto de dados. No contexto do processamento de imagem, esses valores ou informações são representados pelo conjunto de pixels, sendo essenciais para a classificação precisa de cada classe de folhas de erva-mate.

O tempo de extração de características é variável, aumentando conforme mais descritores são combinados. O modelo proposto, apesar de não apresentar seu tempo de extração específico, incorpora uma gama mais abrangente de descritores, incluindo uma abordagem de CNN, sugerindo uma complexidade computacional mais elevada.

O experimento conduzido possibilitou uma análise qualitativa dos descritores no processo de classificação de patologias das folhas de erva-mate, destacando a importância desses elementos na eficácia do modelo proposto.

TABELA 6 - DELINEAMENTO DE EXPERIMENTOS COM BASE EM DESCRITORES

Experimento	Descritores	Instâncias	Características	Tempo
A	Momento Hu	3728	7	23min 3s
B	Textura Harallick	3728	13	55min 42s
C	Histograma de Cores	3728	512	24min 5s
D	Momento Hu + Textura Harallick	3728	20	57min 28s
E	Momento Hu + Histograma de Cores	3728	519	24min 37s
F	Texturas Harallick + Histograma de Cores	3728	525	59min 55s
G	Momento Hu + Textura Harallick + Histograma de Cores	3728	532	1h 40s
Modelo Proposto	Momento Hu + Textura Harallick + Histograma Cores + CNN	3728	532	-

Fonte: Autor (2023).

4.3.3 Resultados de classificação do modelo proposto e outros modelos de aprendizado de máquina

A Tabela 7 apresenta as taxas de classificação das classes das folhas de erva-mate em diferentes experimentos delineados, destacando a abordagem híbrida proposta (Momento Hu + Textura Harallick + Histograma Cores + CNN) e três classificadores baseados em aprendizado de máquina supervisionado (RF, SVM e KNN). O modelo híbrido proposto alcançou uma precisão de 97,8% e acurácia de 97,6%. A taxa de *recall*, indicando a porcentagem de dados classificados corretamente como positivos em relação ao total de positivos reais na amostra, foi de aproximadamente 96,7%. Isso implica que, ao apresentar 100 imagens ao modelo, cerca de 97 imagens serão classificadas corretamente. Ao comparar os três modelos de aprendizado de máquina, o classificador RF demonstrou um desempenho superior, com uma acurácia em torno de 86,4% e precisão de 87%. Em seguida, o classificador SVM obteve uma precisão de 71,7% e acurácia de 70,7%. Por outro lado, o classificador KNN apresentou um desempenho relativamente inferior em relação aos outros dois, com uma precisão de 65,8% e acurácia em torno de 64,6%. No que diz respeito à taxa de recall para as folhas classificadas como positivas em comparação com a quantidade real de positivos na amostra foi de aproximadamente 64,2%. Estes resultados indicam claramente a eficácia da abordagem híbrida e a superioridade do classificador RF neste contexto específico.

TABELA 7 - DESEMPENHO DO MODELO PROPOSTO COM OUTROS CLASSIFICADORES

Modelo	Precisão (%)	Recall (%)	F-Score (%)	Acurácia (%)
Modelo Proposto	97,8	96,7	97,5	97,6
RF	87,0	86,2	86,4	86,4
SVM	71,7	70,2	70,4	70,7
KNN	65,8	64,2	64,5	64,6

Fonte: Autor (2023).

A matriz de confusão gerada pelo modelo híbrido proposto é apresentada na Tabela 8. O modelo apresenta um recall médio de 97%, destacando sua eficácia em identificar de maneira precisa todos os exemplos positivos em relação às diferentes classes. Essa métrica reflete a habilidade do modelo em capturar de maneira abrangente as instâncias corretas, evidenciando um desempenho sólido na categorização das diversas classes de folhas. O cálculo de recall possibilita uma análise detalhada dos resultados, considerando a diversidade das classes foliares. No modelo proposto, as folhas sem patologia foram melhor classificadas, com uma precisão de 97,8%, seguidas pelas folhas com patologias pinta-preta e requeima, ambas com uma taxa de acertos de 97,0%. A alta taxa de classificação ressalta a capacidade dos descritores na extração de características das folhas. Por outro lado, a taxa de recall teve uma leve queda para as folhas com a patologia de fumagina, com taxa de acertos de 96,5%. Este resultado ainda indica uma classificação bastante precisa, evidenciando a eficiência do modelo até mesmo nas classes mais desafiadoras.

TABELA 8 - MATRIZ DE CONFUSÃO DO MODELO HÍBRIDO PROPOSTO

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	228	2	2	1	97,8
Pinta-preta	2	225	4	2	97,0
Fumagina	2	3	224	4	96,5
Requeima	1	3	3	227	97,0
Precisão(%)	97,8	96,6	96,0	97,4	

Fonte: Autor (2023).

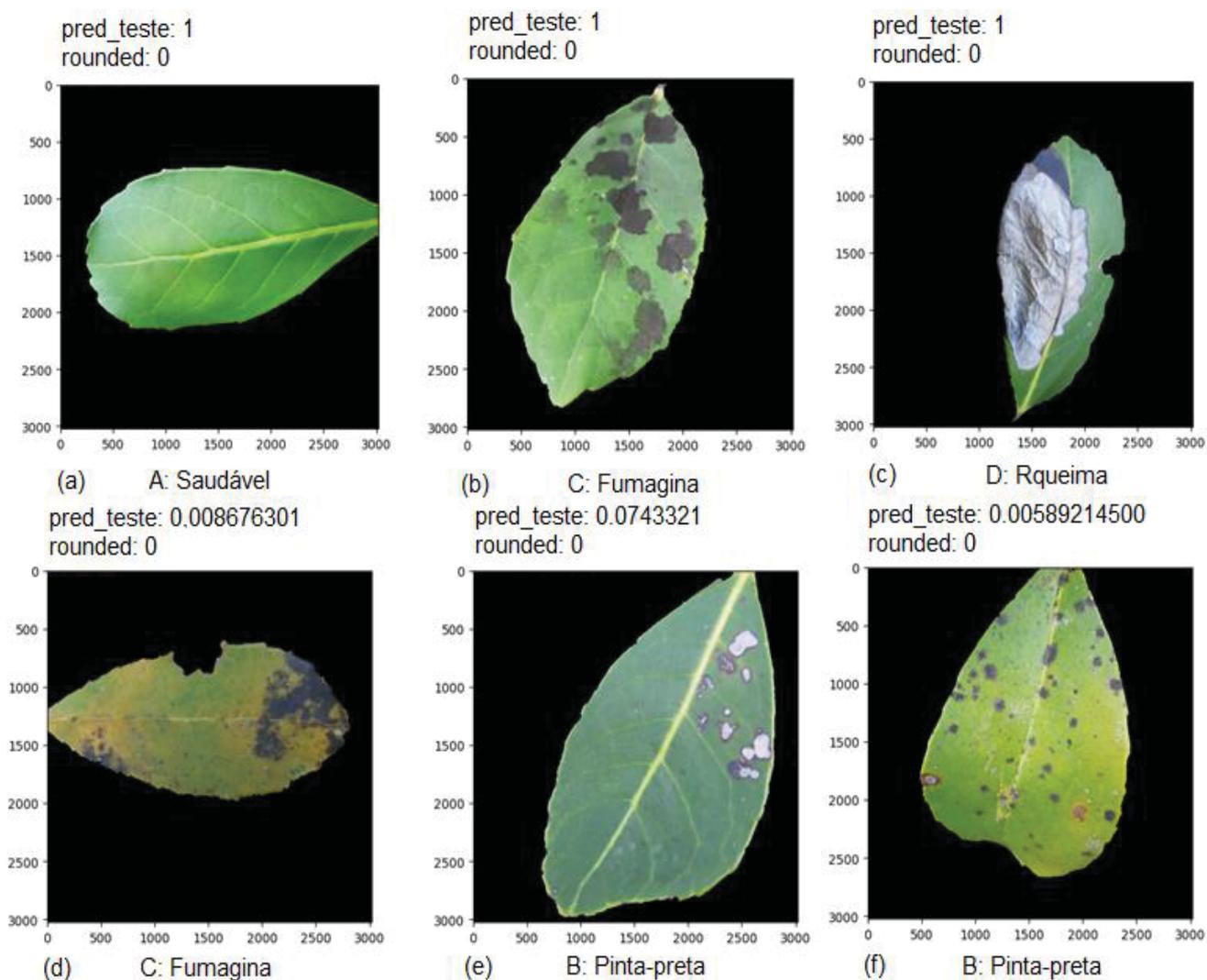
Outro conjunto de dados independente do utilizado no treinamento e validação foi empregado para avaliar a capacidade do modelo proposto na identificação de doenças foliares de erva-mate. Nos testes realizados, os resultados experimentais demonstraram que o modelo híbrido proposto é eficaz e possui um alto nível de precisão de identificação das patologias, alcançando uma acurácia em torno de 95,6%. No entanto, apesar das pontuações consistentes alcançadas pelo modelo,

também foram observadas imprecisões no processo de classificação de folhas de erva-mate.

A Figura 22 ilustra os resultados das previsões, onde as três primeiras imagens de folhas (a), b) e c)) foram identificadas com as pontuações de precisão mais altas, enquanto as três últimas (d), e) e f)) demonstram imprecisões na identificação de doenças foliares. As imprecisões observadas podem estar relacionadas a diversos fatores, incluindo: i) a forma das manchas nas folhas; ii) a textura na imagem; e a cor da área infetada. Esses elementos foram cruciais para os descritores na extração de características para a classificação das folhas. Por exemplo, na imagem d), a folha foi classificada como infectada - a cor da mancha se assemelha à patologia de fumagina, mas foi rotulada manualmente como requeima. Na imagem e), a identificação como pinta-preta pode ter sido afetada por elementos como a forma e a cor, levando à imprecisão na classificação. Na imagem f), também identificada como pinta-preta, a forma da mancha pode ter contribuído para a imprecisão. Além disso, em alguns casos, as imagens das folhas podem apresentar manchas que não são facilmente distinguíveis, aumentando assim a complexidade da tarefa de classificação. Por outro lado, a classificação baseou-se na avaliação do conjunto de dados de entrada a olho humano, ou seja, os dados foram rotulados considerando a mancha visual. Isso pode contribuir para imprecisão de rotulagem de amostra para entrada no modelo. Esses desafios destacam a importância de considerar não apenas a precisão, mas também a complexidade e a variabilidade das características das folhas ao desenvolver modelos de classificação mais robustos e precisos.

A Figura 22 mostra a precisão e imprecisão das previsões de doenças foliares - descrito pelos valores em cada rótulo, onde as imagens a), b) e c) foram identificadas e classificadas com precisão e as imagens d), e) e f), com imprecisão na identificação de doenças no conjunto de dados.

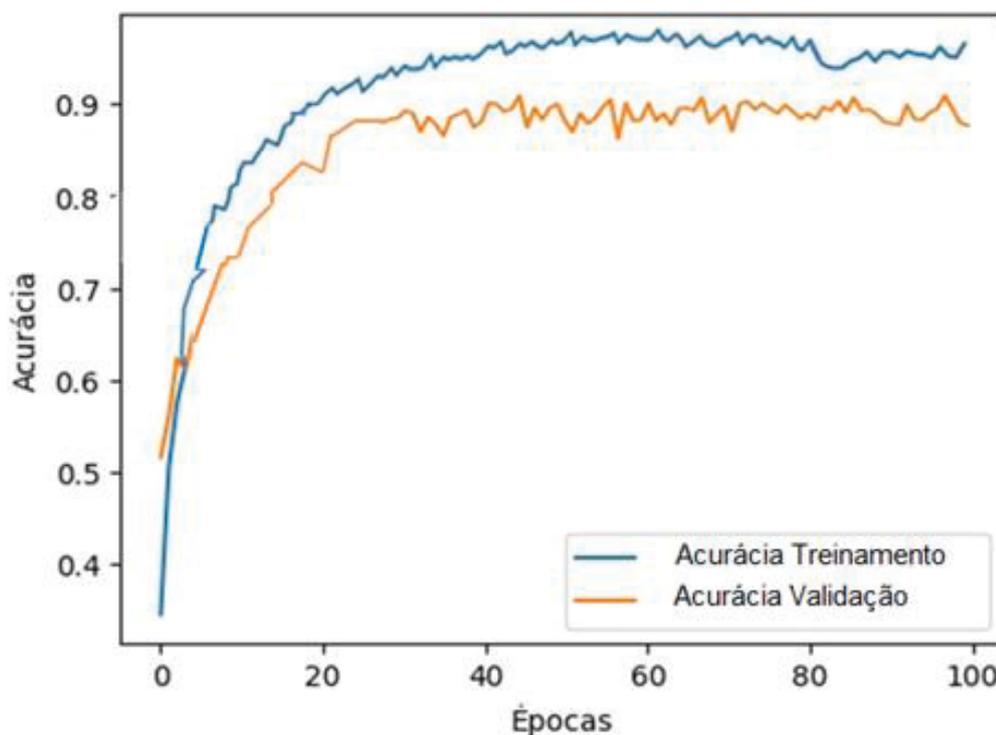
FIGURA 22 - PRECISÃO E IMPRECISÃO DAS PREVISÕES DE DOENÇAS FOLIARES



Fonte: Autor (2023).

Durante as 100 épocas de treinamento e validação, como mostrado na Figura 23, o modelo híbrido proposto alcançou uma pontuação de acurácia de 97,6%. O gráfico revela claramente que à medida que o conjunto de dados de treinamento e o número de épocas aumentam, a acurácia do modelo também aumenta. Esse padrão de melhoria na precisão, com o aumento do treinamento e da exposição a mais dados, destaca a eficácia do modelo híbrido na aprendizagem e na adaptação às complexidades do conjunto de dados, resultando em uma classificação mais precisa das folhas de erva-mate.

FIGURA 23 - INDICADORES DE ACURÁCIA NO TREINAMENTO E VALIDAÇÃO DO MODELO PROPOSTO



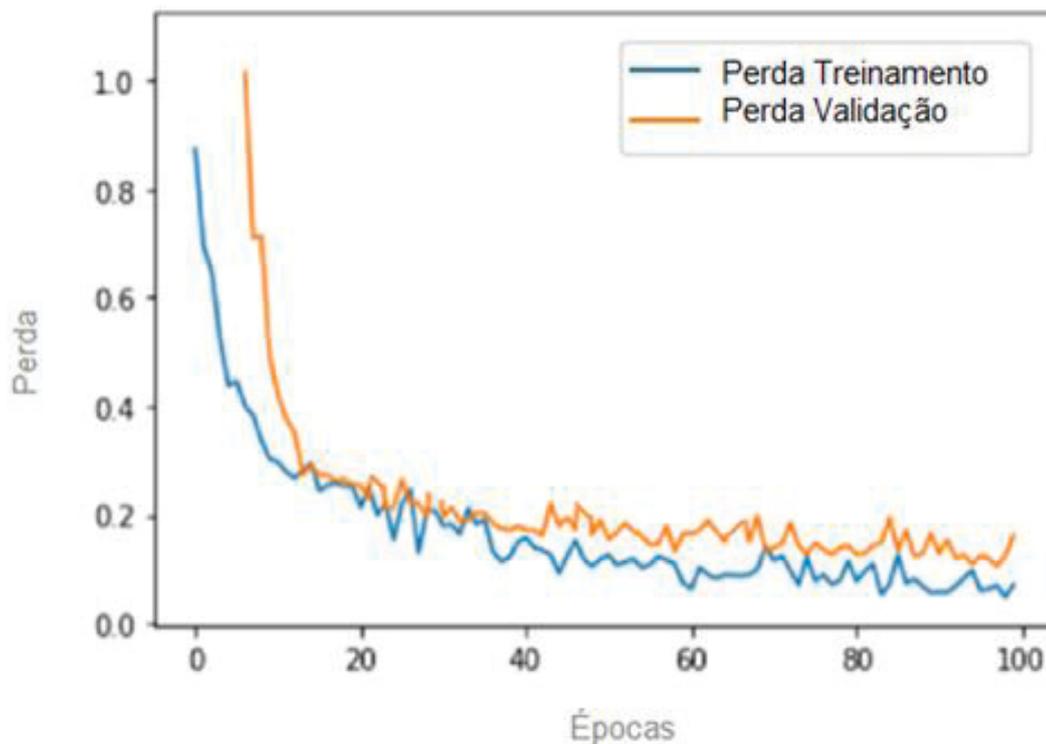
Fonte: Autor (2023).

As curvas de perda de treinamento e validação são apresentadas na Figura 24. No caso do modelo proposto, o valor de perda atinge aproximadamente 0,10%. As perdas são calculadas com base nos erros computados no conjunto de dados de treinamento e são uma medida crítica para avaliar o desempenho do modelo. Essa avaliação é realizada considerando apenas o conjunto de dados de treinamento, ao contrário da métrica de precisão.

A perda representa a soma das discrepâncias entre as previsões do modelo e os valores reais de cada exemplo, tanto no conjunto de dados de treinamento quanto no conjunto de dados de validação. Dessa forma, o valor da perda indica o quão eficaz ou deficiente é o desempenho do modelo após cada iteração de otimização. Uma observação importante é que, se a perda de treinamento for significativamente menor que a perda de validação, isso pode sugerir que o modelo está sofrendo sobreajuste (*overfitting*). Esse fenômeno, conforme observado por Jaiganesh et al. (2020), implica

que o modelo se adaptou excessivamente aos dados de treinamento, tornando-se menos capaz de generalizar para novos dados. Portanto, manter um equilíbrio adequado entre a perda de treinamento e a perda de validação é crucial para garantir a eficácia e a generalização do modelo para dados não vistos anteriormente.

FIGURA 24 - INDICADORES DE PONTUAÇÃO DE PERDA NO TREINAMENTO E VALIDAÇÃO DO MODELO PROPOSTO



Fonte: Autor (2023).

No Experimento G, foi utilizado o classificador de aprendizado de máquina RF combinando os três descritores (Momento Hu + Textura Harallick + Histograma de Cores). As configurações de hiperparâmetros do modelo foram ajustadas para $n_estimators = 100$, indicando que o modelo foi construído com 100 árvores, e $random_state = 9$, uma semente que introduz aleatoriedade no processo, como o embaralhamento dos dados antes de criar as árvores. A análise da matriz de confusão, apresentada na Tabela 9, revela uma taxa média de classificação correta de aproximadamente 85,6%. Essa taxa foi calculada com base no número total de elementos classificados corretamente em todas as classes, em relação ao número total de elementos. As taxas de *recall* foram utilizadas para avaliar os resultados de

acordo com o tipo de folha. Os resultados mostraram que as folhas sem patologia foram classificadas com a maior precisão, atingindo 87,0% de previsões corretas, seguidas pelas folhas com patologia de Requeima, com 86,5%. Por outro lado, a classificação das folhas com patologia de Fumagina apresentou a menor taxa de recall, registrando 84,5%. Essa baixa taxa de *recall* pode ser atribuída à semelhança dos sintomas ou manchas em termos de forma, cor e geometria. Algumas patologias das folhas possuem características estruturais e formas similares, o que pode ter contribuído para a baixa precisão na classificação, como é apresentada na Figura 20, onde *d*) classificada em patologia Fumagina, enquanto foi rotulado manualmente em patologia Requeima, *e*) classificada em patologia Pinta-preta, enquanto a folha está coberta na sua maioria pela patologia Requeima e *f*) ainda classificada em patologia Pinta-preta mas rotulada corretamente em patologia Fumagina.

Os resultados experimentais do teste do modelo também confirmaram a eficácia do modelo, com uma acurácia em torno de 82,7%. Isso sugere que, embora o modelo tenha demonstrado uma boa taxa de classificação, ainda há espaço para melhorias, especialmente na identificação precisa de folhas com patologias semelhantes, com base na análise laboratorial. Esse desafio destaca a complexidade da tarefa de classificação e a importância de considerar características detalhadas e sutis para aprimorar a precisão do modelo em cenários de diagnóstico de doenças foliares.

TABELA 9 - MATRIZ DE CONFUSÃO DE CLASSIFICADOR RANDOM FOREST COMBINANDO OS DESCRITORES

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	201	10	10	10	87,0
Pinta-preta	14	200	13	9	85,0
Fumagina	11	13	196	12	84,5
Requeima	7	10	14	199	86,5
Precisão (%)	86,3	85,8	84,0	86,5	-

Fonte: Autor (2023).

O Experimento G envolveu o uso do classificador de aprendizado de máquina SVM, treinado com um *kernel* = "*linear*" e um parâmetro de regularização $c=1$, para alcançar um equilíbrio moderado entre a suavidade da superfície de decisão e a precisão. Portanto, foi realizada uma análise da matriz de confusão (conforme apresentada na Tabela 10). A precisão do modelo, medida pela porcentagem de elementos corretamente classificados em relação ao total de elementos, atingiu uma média de aproximadamente 71,7%.

Para avaliar o desempenho do modelo de forma mais específica, foram consideradas as taxas de recall, as quais fornecem uma visão detalhada do desempenho em relação aos diferentes tipos de folhas. Os resultados revelaram que o SVM obteve sua melhor performance ao classificar folhas sem qualquer patologia, alcançando uma taxa de recall impressionante de 72,7%. Em seguida, as patologias de Pinta-preta e Requeima foram classificadas com taxas de recall de 72,0% e 71,2%, respectivamente. No entanto, a classificação de folhas com a patologia Fumagina apresentou uma taxa de *recall* relativamente mais baixa. Essa menor precisão pode ser atribuída à similaridade nas características de forma, cor e geometria das manchas que definem essa patologia nas folhas, como mostra a Figura 20 em f), a patologia Fumagina foi classificada em patologia Pinta-preta

Os resultados dos testes do modelo SVM destacaram sua eficácia geral, com uma acurácia em torno de 66,8%. Apesar dos desafios na classificação de determinadas patologias, o modelo demonstrou sua capacidade de identificar corretamente a maioria das classes de folhas, refletindo um desempenho sólido, embora existam oportunidades para aprimoramentos, especialmente nas áreas onde a precisão foi ligeiramente menor.

TABELA 10 - MATRIZ DE CONFUSÃO DE CLASSIFICADOR SUPPORT VECTOR MACHINE COMBINANDO OS DESCRITORES

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	168	19	25	19	72,7
Pinta-preta	25	166	26	14	72,0
Fumagina	20	25	158	20	70,9
Requeima	20	23	24	166	71,2
Precisão (%)	72,1	72,2	68,0	75,8	-

Fonte: Autor (2023).

No modelo KNN, o hiperparâmetro "*n_neighbors*" foi definido como 5, o que implica que o algoritmo considerou os cinco vizinhos mais próximos ao fazer previsões ou classificações. A matriz de confusão resultante do experimento G, mostrada na Tabela 11, apresenta as taxas de *recall*, que tiveram uma média de 64,9%. Esse valor reflete a precisão com que o modelo classificou corretamente as amostras em cada classe, indicando sua habilidade em identificar as classes de maneira precisa. Semelhante aos experimentos anteriores, obteve-se maior sucesso com as folhas sem nenhuma patologia, com uma taxa de acertos de aproximadamente 69,3%, seguidas pelas folhas com patologia de Pinta-preta e Requeima, com taxas de acertos em torno de 68,0% e 61,8%, respectivamente. A taxa de *recall* diminuiu ligeiramente, especialmente para as folhas com patologia Fumagina. Essa redução na precisão pode ser atribuída aos mesmos fatores mencionados nos dois modelos descritos anteriormente.

Neste modelo específico, os resultados experimentais dos testes confirmaram a eficácia do KNN, com uma acurácia em torno de 61,5%. Apesar dos desafios na classificação das folhas, o modelo demonstrou uma capacidade razoável de identificar corretamente várias classes de folhas. No entanto, há espaço para melhorias, especialmente em relação às folhas com patologia Fumagina, onde a precisão foi um pouco menor, refletindo a complexidade da tarefa de classificação nessas condições. A experiência mostrou que a adequada identificação e rotulação da patologia Fumagina requer uma análise visual minuciosa, ou até mesmo análise laboratorial.

Em termos de identificação na imagem digital, é possível que seja necessário usar outros descritores não considerados no presente estudo.

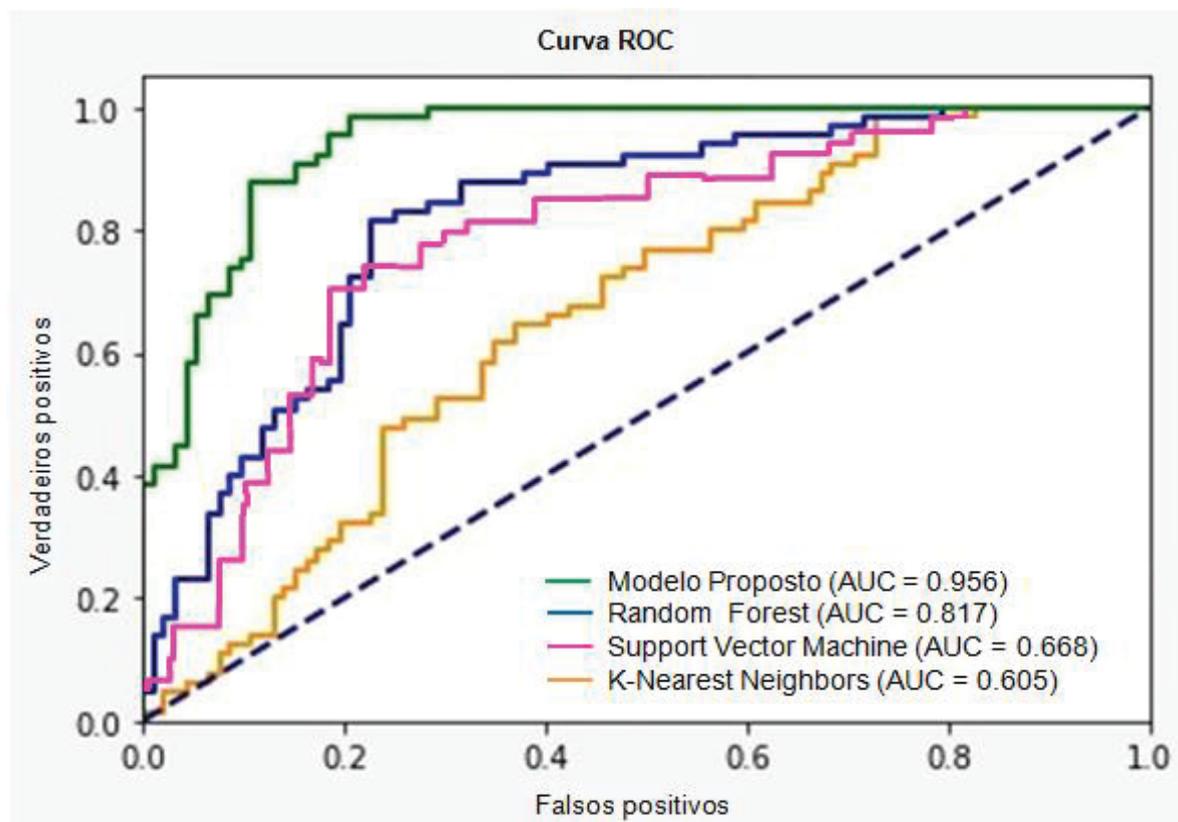
TABELA 11 - MATRIZ DE CONFUSÃO DE CLASSIFICADOR K-NEAREST NEIGHBORS COMBINANDO OS DESCRITORES

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	167	20	36	18	69,3
Pinta-preta	17	134	20	27	68,0
Fumagina	26	33	147	34	61,3
Requeima	23	46	30	160	61,8
Precisão (%)	71,7	57,5	63,0	66,9	-

Fonte: Autor (2023).

Na Figura 25 é feita a análise de curva ROC. A curva ROC é uma ferramenta valiosa para avaliar e comparar o desempenho de diferentes modelos de classificação. Ajuda a escolher um limite de decisão que equilibra sensibilidade e especificidade de acordo com o contexto específico do problema em questão. No caso do estudo, o modelo proposto apresenta melhor desempenho em comparação com os modelos de aprendizado de máquina.

FIGURA 25 - ANÁLISE CURVA ROC-AUC DOS MODELOS



Fonte: Autor (2023).

4.3.4 Avaliação qualitativa dos descritores

Neste estudo, além do modelo proposto, foram conduzidos experimentos para uma avaliação qualitativa dos descritores. A Tabela 12 apresenta o desempenho dos três modelos de aprendizado de máquina em cada experimento. O descritor Histograma de Cores no **Experimento C** alcançou uma pontuação notavelmente alta, demonstrando sua capacidade de classificar o conjunto de dados, que consiste em 4 classes de folhas de erva-mate. Além disso, o **Experimento E**, que combina os descritores Momento Hu e Histograma de Cores, também obteve uma pontuação bastante elevada, seguido pelo **Experimento F**, que combina os descritores Texturas Harallick e Histograma de Cores.

TABELA 12 - ANÁLISE QUALITATIVA DOS MODELOS DE APRENDIZADO DE MÁQUINA NO USO ISOLADO E COMBINAÇÃO DOS DESCRITORES

Experimentos	Métricas	RF	SVM	KNN
A	Acurácia	52,5	46,2	42,7
	Precisão	53,3	48,8	41,8
	Recall	51,5	46,0	41,7
	F-score	52,1	45,6	41,2
B	Acurácia	65,8	60,8	52,8
	Precisão	63,1	58,6	52,1
	Recall	65,2	60,1	52,5
	F-score	63,7	57,9	51,4
C	Acurácia	88,6	75,8	65,0
	Precisão	88,7	76,1	65,5
	Recall	88,3	75,8	64,9
	F-score	88,3	75,5	64,6
D	Acurácia	65,8	58,6	53,6
	Precisão	64,0	57,5	53,8
	Recall	65,9	58,1	53,3
	F-score	64,4	55,9	53,3
E	Acurácia	87,8	76,1	62,8
	Precisão	87,7	76,7	63,2
	Recall	87,6	76,0	62,7
	F-score	87,5	75,7	62,3
F	Acurácia	86,3	70,7	63,5
	Precisão	86,5	72,2	64,3
	Recall	86,3	70,7	63,3
	F-score	86,3	70,4	63,2
G	Acurácia	86,4	70,7	64,6
	Precisão	87,0	71,7	65,8
	Recall	86,2	70,2	64,2
	F-score	86,4	70,4	64,5

Fonte: Autor (2023).

Observa-se na tabela que o uso isolado de descritores individuais, no caso deste conjunto de dados, não resultou em melhorias significativas, ao contrário do desempenho satisfatório alcançado pelo descritor Histograma de Cores. Este fenômeno pode ser atribuído à relevância crucial da cor dos sintomas nas folhas para a quantificação das características. O Histograma de Cores, ao capturar e analisar detalhadamente as variações de cor, demonstrou ser um descritor essencial para distinguir as diferentes classes de folhas com patologias. A capacidade deste descritor

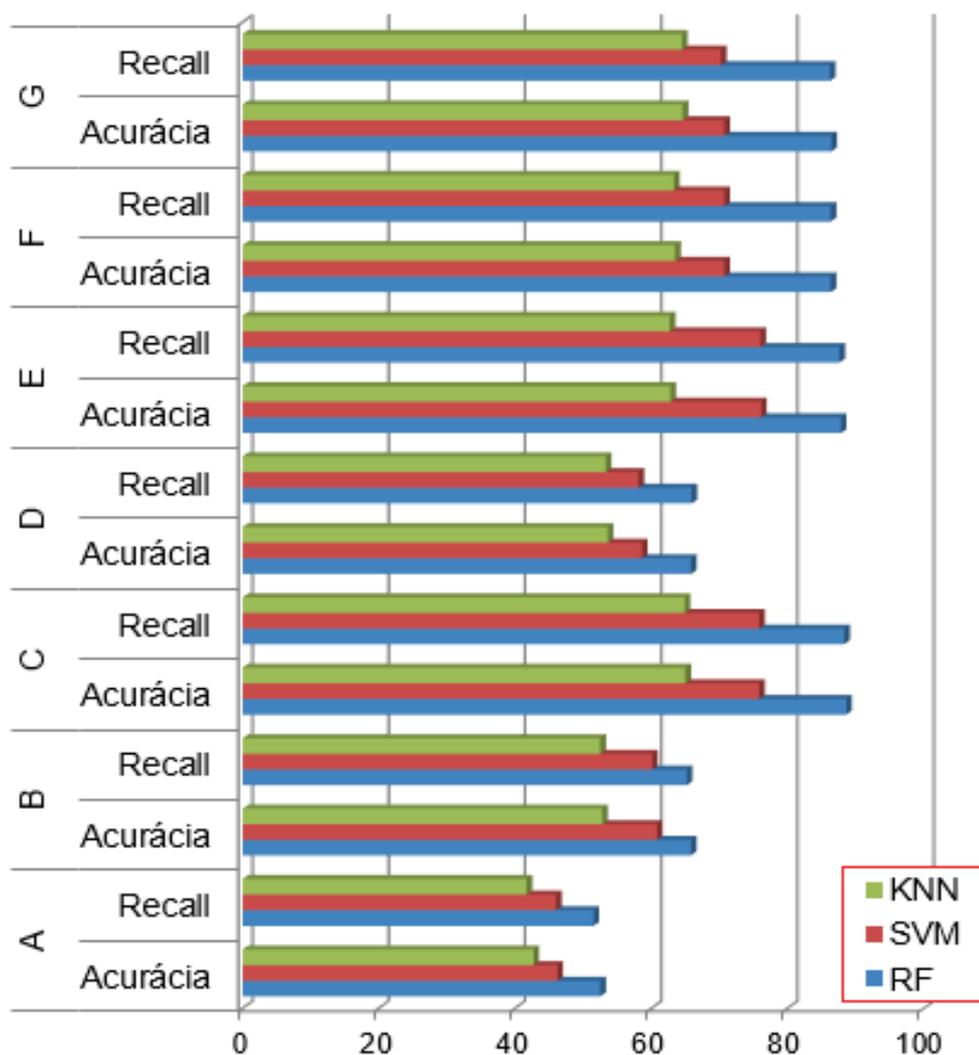
em destacar nuances de cor, considera-se importante na tarefa de classificação, ilustrando assim a sua importância fundamental na extração de características cromáticas para a análise e classificação eficaz de dados complexos como os observados nas folhas de erva-mate.

Estes resultados ressaltam a eficácia desses descritores específicos na classificação precisa das diferentes classes de folhas, fornecendo informação valiosa para a escolha dos melhores descritores em tarefas similares no futuro.

A Figura 26 mostra o desempenho dos modelos em cada experimento, destacando tanto a acurácia global quanto a taxa de recall. A taxa de recall representa a proporção de elementos corretamente classificados em todas as classes em relação ao número total de elementos, fornecendo uma medida essencial da precisão do modelo.

Nos experimentos C, E, F e G, os modelos Random Forest (RF) e, em seguida, o Support Vector Machine (SVM), apresentaram os melhores resultados. Isso evidencia a eficácia dos descritores na extração de características para a classificação do conjunto de dados, especialmente quando se utilizam os descritores Histograma de Cores, a combinação de Momento Hu e Histograma de Cores, assim como Texturas Harallick e Histograma de Cores. A habilidade desses modelos em capturar nuances específicas das imagens, permitindo uma diferenciação precisa entre as classes de folhas, destaca a importância desses descritores na tarefa de classificação. Esses resultados reforçam a relevância desses métodos para análises futuras e aplicações práticas no contexto de diagnóstico de doenças foliares em plantas como a erva-mate.

FIGURA 26 - DESEMPENHO DOS MODELOS NOS DIFERENTES EXPERIMENTOS



Fonte: Autor (2023).

Na Tabela 13, estão apresentados os resultados obtidos por meio da técnica de validação cruzada realizada 20 vezes. Ao considerar as pontuações da curva ROC-AUC, fica evidente que o modelo RF alcançou a maior precisão na classificação das doenças foliares de erva-mate, seguido do modelo SVM, destacando a eficácia desses modelos na tarefa de discriminação e diagnóstico preciso das doenças foliares em plantas.

TABELA 13 - ANÁLISE DE CURVA ROC-AUC DOS MODELOS EM VALIDAÇÃO CRUZADA

Experimento	Descritores	RF		SVM		KNN	
		ROC-AUC	STD	ROC-AUC	STD	ROC-AUC	STD
A	MH	0,53	0,041	0,47	0,038	0,44	0,035
B	TH	0,65	0,040	0,61	0,049	0,53	0,042
C	HC	0,87	0,046	0,76	0,046	0,65	0,044
D	MH, TH	0,65	0,056	0,58	0,041	0,54	0,046
E	MH, HC	0,87	0,052	0,76	0,043	0,63	0,037
F	MH, TH, HC	0,87	0,052	0,71	0,041	0,64	0,057
G	MH, TH, HC, CNN	0,87	0,046	0,72	0,045	0,66	0,062

* Momento Hu (MH); Textura Harallick (TH); Histograma Cores (HC).

Fonte: Autor (2023).

4.3.5 Tempo de processamento

Além de avaliar os modelos com base na precisão e outras métricas relacionadas combinando os três descritores, também foi analisado o tempo de treinamento em diferentes modelos conforme a Tabela 14. Nos experimentos realizados, observou-se que à medida que aumentava o número de épocas e neurônios no aprendizado profundo, o tempo de treinamento do modelo aumentava consideravelmente. Para alcançar uma precisão de 97,8% em 100 épocas, o modelo demandou aproximadamente 3240 segundos. Em contraste, o tempo de treinamento dos modelos de aprendizado de máquina foi significativamente menor, com o RF levando 362 segundos e SVM e KNN demandando 550 segundos e 548 segundos, respectivamente.

TABELA 14 - TEMPO DE TREINAMENTO COMBINANDO OS TRÊS DESCRITORES

Modelo	Tempo (s)
Modelo Proposto	3240
RF	362
SVM	550
KNN	548

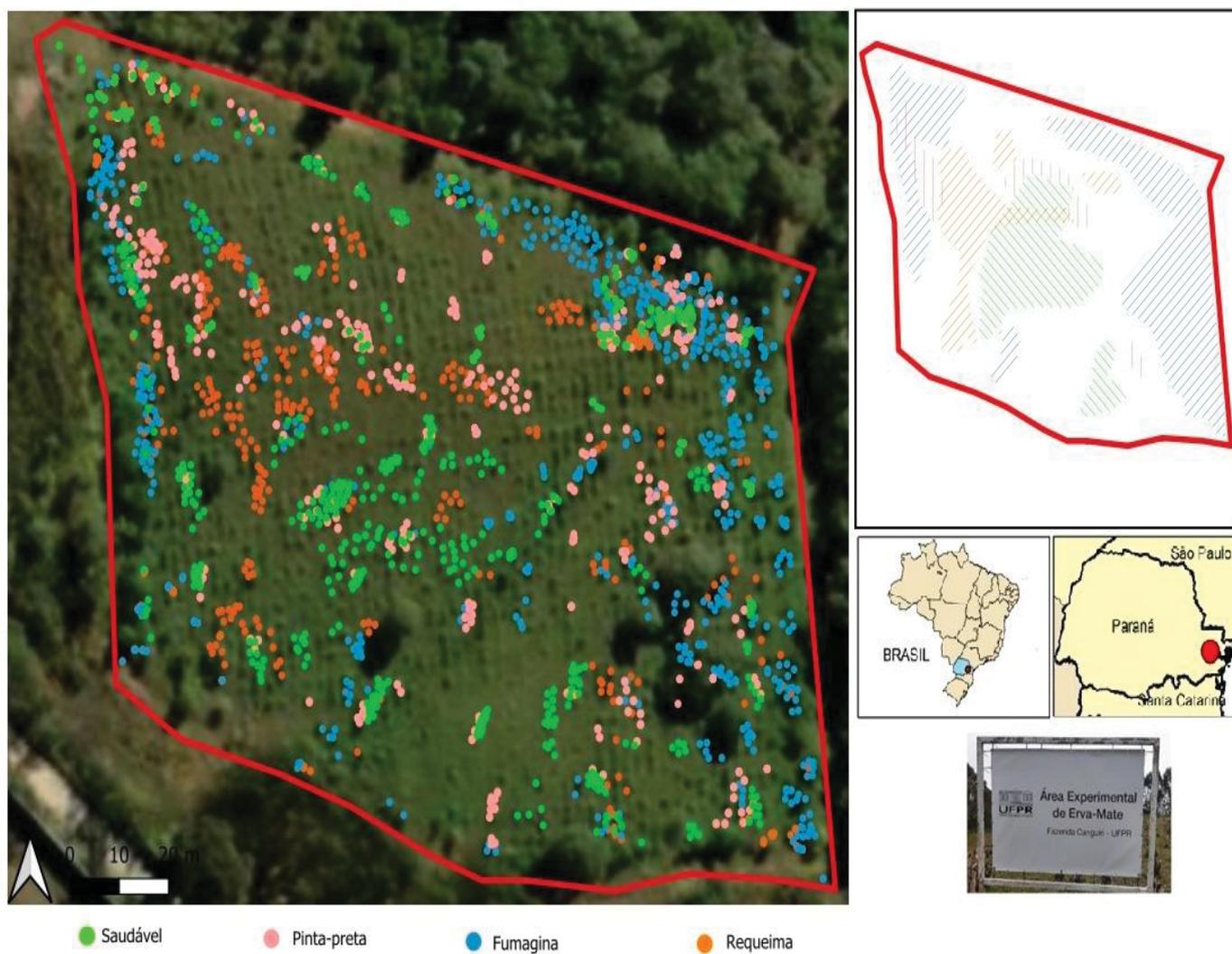
Fonte: Autor (2023).

4.4 MAPEAMENTO DA OCORRÊNCIA DE PATOLOGIAS NA ÁREA DE ESTUDO

A Figura 27 ilustra o mapa de ocorrência de patologias com base nas informações posicionais fornecidas pela câmera. De acordo com o mapa, as folhas sem nenhuma patologia estão predominantemente localizadas na parte central do plantio. A patologia fumagina tem sua incidência mais alta nas bordas do plantio, onde há mais sombras devido à cobertura das árvores. A pinta-preta está distribuída por quase toda a área de plantio, com maior incidência no noroeste. As folhas com requeima concentram-se principalmente no noroeste e na parte central, provavelmente devido à maior exposição ao sol nessas áreas.

Os resultados produzidos mostram que, combinando as técnicas de segmentação e classificação de imagens com métodos de posicionamento permite mapear a ocorrência de doenças e descrever seus padrões espaciais. Chama a atenção a distribuição da patologia fumagina, que ocorre principalmente nas bordas do plantio, contrastando com a alta ocorrência de folhas saudias no interior do conjunto. Isto pode mostrar algum efeito da exposição ou proteção das plantas em função da distribuição espacial das folhas.

FIGURA 27 - MAPA OCORRÊNCIA DE PATOLOGIAS NA ÁREA EXPERIMENTAL DE ERVA-MATE



Fonte: Autor (2023).

5 DISCUSSÃO DOS RESULTADOS

5.1 DETECÇÃO E SEGMENTAÇÃO DA FOLHA DE ERVA-MATE

As imagens das folhas de erva-mate foram rotuladas manualmente com pixels de destino e de fundo, e esses dados verdadeiros foram utilizados no experimento. Quatro modelos de segmentação foram treinados e avaliados em termos de eficácia e desempenho, levando em consideração suas características e abordagens distintas. Os resultados experimentais indicaram que é possível segmentar corretamente as imagens com fundo complexo utilizando métodos de aprendizado profundo de última geração, como YOLOv8 proposto neste estudo. O teste revelou que o algoritmo é capaz de segmentar as folhas de cultura de erva-mate com alto grau de certeza, obtendo um índice mAP de aproximadamente 99,1% para uma média de precisão com uma sobreposição de 50%. Para uma média de precisão com uma sobreposição de 95% obteve-se índice mAP de 97,7%, ainda considerada satisfatório para a tarefa de detecção e segmentação. A diferença entre mAP50 (99,1%) e mAP50-95 (97,7%) na tarefa de detecção e segmentação de folhas indica como o modelo se comporta em relação à sobreposição de caixas delimitadoras em diferentes faixas. A diferença entre mAP50 e mAP50-95 indica que o modelo tem uma excelente capacidade de detecção mesmo com uma sobreposição mínima de 50%, mas essa precisão diminuiu um pouco quando são considerados requisitos mais rigorosos de sobreposição na faixa de 50% a 95%.

Comparando os valores de ME médio para cada algoritmo, obteve-se os seguintes resultados: YOLOv8 (1,28%), Mask-RCNN (4,28%), U-Net (8,13%) e GrabCut (24,65%). O YOLOv8 demonstrou um Erro Médio de 1,28%, indicando uma precisão relativamente alta na tarefa de segmentação. O baixo valor de ME sugere que o YOLOv8 é eficaz na predição de caixas delimitadoras, o que é fundamental em tarefas de segmentação de objetos.

Os resultados experimentais confirmaram que o modelo proposto supera os métodos comparativos descritos na Tabela 15, que foram testados em diferentes pesquisas sobre a tarefa de segmentação.

TABELA 15 - DESCRIÇÃO DE DIFERENTES ESTUDOS COMPARATIVOS AO MODELO PROPOSTO

Autor	Banco Dados	Método	Acurácia (%)	IOU (%)	mAP (%)
Yang et al. (2020)	-	Mask R-CNN	-	-	98,0
Storey et al. (2022)	Plant Pathology Challenge 2020	Mask R-CNN [ResNet-50, MobileNetV3-Large e MobileNetV3-Large-Mobile]	80,5	-	-
Wang et al. (2021)		DeepLabv3+	-	90,3	
Patidar and Vianny (2020)	CVPPP-2017	U-Net	-	92,7	-
Chiu et al. (2020)	UAV RGB Images	U-Net and DeepLabv3+	-	93,20	-
Jia et al. (2022)		UNet	-	83,58	89,18
Hasan et al. 2022	RoCole	GrabCut		95,0	
Ünver and Ayan (2019)	PH2 e ISBI 2017	YOLO-GrabCut	PH2 (94,4) ISBI2017 (96)	PH2(90) ISBI2017 (86)	-
Wu et al. (2022)	Específico	GrabCut-Mask R-CNN	-	-	95,0
Chen et ai. (2021)	Específico	U-Net	-	91,2	-
		DeepLabv3+		89,5	
		BLSNet		95,6	
Lin et al. (2019)	Específico	U-Net	Precisão 83,45	72,11	-
3 Método comparativos	Específico	GrabCut	89,29		-
	Específico	U-Net	95,30	-	-
	Específico	Mask R-CNN	98,9	97,4	-
Método Proposto	Específico	YOLOv8	99,9	99,1	40,3

Fonte: Autor (2023).

Destacam-se os estudos de Hasan et al. (2022), Lin et al. (2019), Patidar e Vianny (2020), Wang et al. (2021), Chiu et al. (2020), Chen et al. (2021), Jia et al. (2022) e Storey et al. (2022). Alguns desses estudos alcançaram resultados concorrentes ao modelo proposto. Por exemplo, Yang et al. (2020) utilizou o algoritmo Mask R-CNN e obteve um índice mAP de 98,0%. Wu et al. (2022), ao combinar o algoritmo tradicional GrabCut com o modelo de aprendizado profundo Mask R-CNN,

alcançou um mAP de 95,0%. Ünver e Ayan (2019), ao combinar YOLO-GrabCut, alcançou uma acurácia de 96,0%. Além disso, no estudo de Liu et al. (2022), que visava contornar os efeitos de segmentação inadequada em fundos complexos e dados em tempo real, foram obtidos bons resultados utilizando o YOLOv5 para detecção de roupas, o algoritmo KMeans para obter a área aproximada do primeiro plano da roupa, e o GrabCut para segmentar a imagem.

5.2 CLASSIFICAÇÃO DE PATOLOGIAS DE ERVA-MATE

O objetivo principal da pesquisa foi desenvolver modelos eficientes na identificação e classificação de doenças das plantas, especialmente focando em folhas de erva-mate, por meio de imagens em RGB. Uma revisão da literatura foi conduzida para entender as limitações e as lacunas de pesquisa existentes. Esta revisão destacou a necessidade premente de modelos capazes de generalizar melhor o conjunto de dados. Na tentativa de abordar esse desafio, foram exploradas soluções envolvendo abordagens híbridas que combinam descritores artesanais de extração de características com técnicas de aprendizado profundo. Além disso, o modelo desenvolvido foi comparado com abordagens convencionais baseadas em algoritmos como Random Forest (RF), Support Vector Machine (SVM) e K-Nearest Neighbors (KNN). Durante essa avaliação comparativa, métricas fundamentais como acurácia, precisão, recall e F-score foram empregadas. Esse processo de comparação foi essencial para avaliar o desempenho relativo do modelo proposto em relação às técnicas estabelecidas, proporcionando entendimento valioso sobre a eficácia e a robustez do método desenvolvido.

Os artigos selecionados para discussão foram organizados em três categorias principais. Na primeira categoria, estão os estudos que utilizam métodos de aprendizado profundo, assim como transferência de aprendizado, para a classificação de patologias. Na segunda categoria, estão os estudos que empregam descritores para extração de características, combinados com classificadores convencionais de

aprendizado de máquina. Por fim, na terceira categoria, encontram-se estudos que exploram a combinação de descritores artesanais com técnicas de aprendizado profundo para o reconhecimento de doenças em plantas.

O modelo híbrido proposto alcançou uma acurácia em torno de 97,6%, precisão de 97,8% e recall de 96,7%. Estas taxas são altamente satisfatórias para um uso prático e lógico. Os resultados do experimento indicam que o modelo desenvolvido apresentou um desempenho satisfatório na tarefa de reconhecimento de doenças foliares de plantas com base em imagens, concorrendo com outros estudos que utilizaram os mesmos descritores de extração de características. Em particular, destaca-se o estudo de Joseph et al. (2022) na área de saúde para detecção e classificação de câncer de mama que, ao empregar os descritores Momento Hu, Textura Haralick e Histograma de Cores, alcançou melhores resultados em diferentes épocas: 97,87% para 40x, 97,60% para 100x, 96,10% para 200x e 96,84% para 400x de ampliação. Além disso, o estudo de Shetty et al. (2022) ainda na área de saúde, também obteve um desempenho satisfatório, atingindo cerca de 95,18% ao utilizar os mesmos descritores, embora ligeiramente inferior em comparação com os resultados encontrados neste estudo. Estes dados destacam a robustez do modelo proposto, evidenciando sua eficácia na identificação precisa de doenças foliares com base nas imagens capturadas.

Os resultados obtidos pelo modelo desenvolvido mostraram uma vantagem comparativa em relação a estudos que empregaram modelos baseados em redes neurais profundas e técnicas de aprendizagem por transferência. Este melhor desempenho foi notada ao utilizar o conjunto de dados específico em questão, conforme evidenciado nos estudos de Kawasaki et al. (2015), Fujita et al. (2016), Ramcharan et al. (2017), Ma et al. (2018), Sharif et al. (2018), Picon et al. (2019), Selvaraj et al. (2019), Hu et al. (2019) e Zhang et al. (2019).

Além disso, uma melhoria ligeira do modelo desenvolvido foi observada ao compará-lo com os resultados do estudo de Singh et al. (2022) que foram utilizados descritores como momentos de cor, HOG e GLCM para extração de características, combinados com métodos profundos como LeNet, ShuffleNet, AlexNet, EffNet e

MobileNet para a classificação. O modelo MobileNet, otimizado com SVM, alcançou uma acurácia de 92,2%, enquanto o MobileNet, otimizado com Random Forest (RF), atingiu uma acurácia de 96,1%. Olhando os resultados encontrados por Singh et al. (2022), destaca-se que, embora os modelos de aprendizagem por transferência tenham sido considerados melhores em várias classes de doenças nas quais foram treinados, o modelo desenvolvido na presente pesquisa superou as abordagens desenvolvidas pelo Singh et al. (2022), ressaltando sua eficácia na identificação precisa das doenças foliares específicas estudadas.

Nos estudos de Gunarathna e Rathmayaa (2020) e Sharma et al. (2020), que empregaram redes neurais convolucionais tanto como extratores quanto como classificadores para identificação de doenças em plantas. Esses estudos alcançaram uma melhor taxa de acurácia de 98,0%. Além disso, o estudo de Radhakrishnan et al. (2020) desenvolveu um modelo híbrido, integrando uma CNN como extratora de características e um classificador convencional de aprendizado de máquina SVM para a detecção de patologias, alcançando uma taxa de acurácia em torno de 96,8%. Os resultados desses autores demonstram a eficácia das abordagens baseadas em CNN na identificação precisa de doenças em plantas. Os resultados alcançados dos autores acima mencionados alinham-se de perto com os encontrados na presente pesquisa, corroborando a qualidade e precisão do modelo proposto. Isso confirma a competitividade e confiabilidade do método desenvolvido neste estudo em relação a outras abordagens modernas, enfatizando sua relevância no campo do diagnóstico de doenças foliares em plantas.

No presente estudo, os classificadores de aprendizado de máquina convencionais, quando combinados com os três descritores artesanais para extração de características, alcançaram as seguintes taxas de acurácia: RF (86,4%), SVM (70,7%) e KNN (64,6%). Em comparação com os resultados do estudo de Basavaiah & Anthony (2020), que utilizaram os mesmos descritores, a acurácia do modelo RF foi relativamente melhor, com acurácia de 94%. No estudo de Shetty et al. (2022) realizado na área de saúde para detecção e classificação de lesões da pele, utilizando os mesmos descritores, a acurácia do modelo RF foi quase a mesma, enquanto as

pontuações dos modelos SVM e KNN foram relativamente inferiores às encontradas neste estudo.

Resultados semelhantes foram obtidos em estudos anteriores. Por exemplo, no estudo de Suresha et al. (2017), que usou características geométricas para extração de características, o modelo KNN obteve uma acurácia em torno de 76,59%. No estudo de Ramesh et al. (2018), que utilizou o histograma de gradientes orientados (HOG), Momentos de Hu, Textura Haralick e Histograma de cor, o classificador RF obteve uma acurácia de 70,14%. Em outro estudo de Nair et al. (2019) que empregou técnicas de extração de características baseadas em cor, o classificador SVM alcançou uma acurácia de 80,0%. Sehgal (2019), usando descritores de cor e textura, obteve acurácias nos classificadores SVM (72,92%), RF (71,88%), NB (70,57%) e DT (64,0%). Além disso, no estudo de Azim et al. (2021), que utilizou descritores de cor, forma e textura (GLCM), as taxas nos classificadores XGBoost e SVM com kernel RBF foram 86,58% e 81,67%, respectivamente. Apesar de diferentes estudos utilizarem variedade de descritores para extração de características, os resultados encontrados neste estudo não diferem significativamente dos resultados dessas pesquisas anteriores, o que valida a eficácia dos métodos utilizados nesta análise.

Outra constatação considerável observada ao comparar os resultados deste estudo com a literatura existente é que, em alguns casos, há uma pequena diferença em termos de precisão e acurácia na classificação de patologias quando são utilizados conjuntos de dados específicos em comparação com conjuntos de dados públicos. Esta discrepância é evidente nos estudos que empregaram modelos de aprendizado de máquina, como Gharge & Singh (2016), Pujari et al. (2016), Zhang et al. (2017), Zhang et al. (2018), Kaur et al. (2018), Sharif et al. (2018), Khan et al. (2018), Hlaing et al. (2018), Dhingra et al. (2019), Vamsidhar et al. (2019), Jaisakthi (2019), Hossain et al. (2019), Doh et al. (2019), Kurmi et al. (2020), Rao et al. (2020), Basavaiah e Anthony (2020) e Kaur et al. (2021). Um cenário semelhante pode ser observado em estudos que utilizaram métodos de aprendizagem profunda, como os conduzidos por Mohanty et al. (2016), Amara et al. (2017), Brahim et al. (2017), Ferentinos (2018), Too et al. (2019), Abdu et al. (2020), Afifi et al. (2020), Chen et al. (2022), Saberi Anari (2022) e Hassan and Maji (2022). No entanto, algumas dessas discrepâncias podem

estar relacionadas ao tamanho da amostra de treinamento, incluindo as condições em que as imagens foram capturadas, entre ambiente controlado ou não. Esses fatores podem desempenhar um papel significativo nas diferenças observadas nos resultados entre estudos específicos e públicos.

Os algoritmos de aprendizado profundo, incluindo as CNN's, possuem a capacidade intrínseca de extrair automaticamente recursos espaciais complexos e de alta dimensão a partir de dados brutos durante o processo de treinamento (Magsi et al., 2020; Hassan et al., 2021; Kumar et al., 2022; Shi et al., 2023). Apesar dessa vantagem notável, o uso de CNNs enfrenta uma barreira significativa. Muitas vezes é frequentemente desafiador estimar os parâmetros do modelo de forma ideal para a extração de características espaciais sofisticadas quando se trabalha com conjuntos de dados de treinamento pequenos e semelhantes (Hu et al., 2015; Kamilaris e Prenafeta-Boldú, 2018). Em situações em que as imagens são coletadas em ambientes não controlados, essa complexidade é ampliada, pois a presença de interferências e variações imprevisíveis pode comprometer substancialmente o desempenho dos algoritmos de detecção e classificação.

6 CONSIDERAÇÕES FINAIS

A inteligência artificial e as técnicas de fotogrametria desempenham um papel essencial no avanço das pesquisas tecnológicas no setor agrícola. Este estudo propõe o desenvolvimento de uma metodologia para a detecção e classificação de patologias nas folhas da cultura de erva-mate, fazendo uso de métodos de aprendizado de máquina e descritores espectrais e de textura derivados das imagens.

No estudo foi implementado o algoritmo YOLOv8 para detecção e segmentação de folhas de erva-mate em condições reais de campo e mostrou ser significativamente rápido, tornando-o uma excelente escolha para a tarefa.

O uso deste modelo alcançou resultados robustos de segmentação de instâncias de várias folhas de erva-mate, atingindo um índice mAP de 99,1% e um ME de 1,28%. Esse desempenho foi obtido mesmo em situações de forte oclusão, sobreposição e interseções entre diferentes folhas e outras plantas, bem como em imagens com iluminação variável e fundos complexos. O método proposto superou outros métodos testados no estudo, como o Mask R-CNN com um índice mAP de 97,4% e ME de 4,28%, o U-Net com 95,30% e ME de 8,13%, e o GrabCut, que apresentou o maior ME de 24,65%.

O estudo apresentou técnicas de extração de características a serem utilizadas no desenvolvimento de sistema automático para a detecção e classificação de doenças de plantas. A combinação de vetores de características de forma, textura e cor proporcionou uma precisão satisfatória em comparação com o uso de apenas um descritor, especialmente em situações em que as imagens são capturadas em ambientes não controlados. Extrair e escolher as características apropriadas é crucial para um sistema eficiente de detecção e classificação de doenças em folhas, principalmente que há semelhanças visuais em manchas visíveis das patologias. No estudo, observou-se que o descritor Histograma de Cores, assim como a combinação dos descritores Momento Hu e Histograma de Cores, juntamente com Texturas Harallick e Histograma de Cores, alcançaram pontuações relativamente altas. Isso

destaca a importância desses descritores na extração de informações distintivas das imagens, contribuindo para um sistema de classificação preciso e confiável de doenças em plantas.

A principal contribuição do estudo reside na apresentação de um modelo que integra descritores artesanais de extração de características com técnicas de aprendizagem profunda para a classificação de patologias com base em imagens de folhas de erva-mate. Os resultados experimentais evidenciaram que o modelo híbrido desenvolvido possui um desempenho satisfatório, com uma acurácia em torno de 97,6%, precisão de 97,8% e recall de 96,7%. Notavelmente, o modelo conseguiu classificar com precisão as folhas sem nenhuma patologia, alcançando 97,8% de acertos, seguido pelas folhas com patologias pinta-preta e requeima, ambas com uma taxa de acerto de 97,0%. Essas taxas são consideradas satisfatórias para aplicação prática e refletem a eficácia do modelo na identificação precisa das diferentes classes de doenças das quais foi treinado.

Para fins de comparação, foram desenvolvidos três modelos de aprendizado de máquina convencionais para a mesma tarefa. O classificador RF obteve melhor desempenho, alcançando uma acurácia de aproximadamente 86,4%, seguido pelo modelo SVM com 70,7%. Em ambos os modelos, as folhas sem nenhuma patologia foram classificadas com maior precisão, seguidas pelas folhas com patologias de Pinta-preta e Requeima.

Com a integração de imagens geocodificadas, como as fotografadas no plantio de erva-mate, foi possível mapear a frequência de ocorrência de patologias. Esse processo robustece a capacidade de monitorar doenças e pragas ao longo do tempo e espaço, proporcionando informações valiosas para a pesquisa e gestão agrícola. Este progresso é vital para o aprimoramento contínuo das estratégias de controle de doenças e pragas, permitindo ações mais eficazes na preservação das plantações e na sustentabilidade dos cultivos.

6.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Os resultados apresentados na presente pesquisa podem ser melhorados, seja refinando o modelo proposto ou explorando modelos alternativos, primeiro, testando outras técnicas de segmentação e de pré-processamento que possam discriminar melhor as manchas de doenças nas imagens foliares; segundo avaliando diferentes descritores de extração de características e; testando combinações de múltiplos descritores para detecção de patologias em plantas.

Deve-se verificar a viabilidade de se explorar esta abordagem para o desenvolvimento de aplicativos de smartphone para detecção e classificação em ambientes reais de cultivo para suporte aos produtores rurais

Como um dos grandes desafios para o uso de métodos de aprendizado de máquina é dispor de um banco de dados com amostra suficiente para os experimentos, pode-se explorar o conjunto de dados da presente pesquisa para trabalhos futuros.

Outras linhas de pesquisas podem ser direcionadas para a identificação de múltiplas doenças e a avaliação de gravidade das mesmas, pois muitas vezes, as plantas são infectadas por várias doenças simultaneamente, tornando a identificação dessas múltiplas doenças um desafio complexo em comparação com o diagnóstico de uma única doença na folha. Além disso, enquanto a maior parte das pesquisas existentes se concentra na detecção da presença de patologias em plantas, há uma escassez de modelos que visam estimar a gravidade dessas doenças. Essa informação é de extrema importância para os agricultores, destacando a necessidade de investigações mais aprofundadas nesse campo específico.

7 REFERÊNCIAS

- Abdu, A. M., Mokji, M. M., & Sheikh, U. U. (2020). Automatic vegetable disease identification approach using individual lesion features. *Computers and Electronics in Agriculture*, 176, 105660.
- Abdullah, M. N., Shukran, M. A. M., Isa, M. R. M., Ahmad, N. S. M., Khairuddin, M. A., Yunus, M. S. F. M., & Ahmad, F. (2021). Colour features extraction techniques and approaches for content-based image retrieval (CBIR) system. *Journal of Materials Science and Chemical Engineering*, 9(7), 29-34.
- Afifi, A., Alhumam, A., & Abdelwahab, A. (2020). Convolutional neural network for automatic identification of plant diseases with limited data. *Plants*, 10(1), 28.
- Ahmed, I., & Yadav, P. K. (2023). Plant disease detection using machine learning approaches. *Expert Systems*, 40(5), e13136.
- Ahmed, K., Shahidi, T. R., Alam, S. M. I., & Momen, S. (2019, December). Rice leaf disease detection using machine learning techniques. In *2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)* (pp. 1-5). IEEE.
- Ali, S., Hassan, M., Kim, J. Y., Farid, M. I., Sanaulah, M., & Mufti, H. (2022). FF-PCA-LDA: intelligent feature fusion based PCA-LDA classification system for plant leaf diseases. *Applied Sciences*, 12(7), 3514.
- Almakady, Y., Mahmoodi, S., & Bennett, M. (2020). Adaptive volumetric texture segmentation based on Gaussian Markov random fields features. *Pattern Recognition Letters*, 140, 101-108.
- Amara, J., Bouaziz, B., & Algergawy, A. (2017). A deep learning-based approach for banana leaf diseases classification. *Datenbanksysteme für Business, Technologie und Web (BTW 2017)-Workshopband*.
- Andrearczyk, V., & Whelan, P. F. (2016). Using filter banks in convolutional neural networks for texture classification. *Pattern Recognition Letters*, 84, 63-69.
- Andrearczyk, V., & Whelan, P. F. (2017). Deep learning in texture analysis and its application to tissue image classification. In *Biomedical texture analysis* (pp. 95-129). Academic Press.
- Andrearczyk, V., & Whelan, P. F. (2018). Convolutional neural network on three orthogonal planes for dynamic texture classification. *Pattern Recognition*, 76, 36-49.

- Anton, A., Rustad, S., Shidik, G. F., & Syukur, A. (2021). Classification of Tomato Plant Diseases Through Leaf Using Gray-Level Co-occurrence Matrix and Color Moment with Convolutional Neural Network Methods. In *Smart Trends in Computing and Communications: Proceedings of SmartCom 2020* (pp. 291-299). Springer Singapore.
- Anwer, A., & Singh, G. (2019). Geo-spatial technology for plant disease and insect pest management. *Bulletin of Environment, Pharmacology and Life Sciences*, 8(12), 1-12.
- Aruraj, A., Alex, A., Subathra, M. S. P., Sairamya, N. J., George, S. T., & Edwards, S. V. (2019, March). Detection and classification of diseases of banana plant using local binary pattern and support vector machine. In *2019 2nd International Conference on Signal Processing and Communication (ICSPC)* (pp. 231-235). IEEE.
- Avdeenko, T. V., & Makarova, E. S. (2017). Integration of case-based and rule-based reasoning through fuzzy inference in decision support systems. *Procedia Computer Science*, 103, 447-453.
- Azim, M. A., Islam, M. K., Rahman, M. M., & Jahan, F. (2021). An effective feature extraction method for rice leaf disease classification. *Telkomnika (Telecommunication Computing Electronics and Control)*, 19(2), 463-470.
- Bagri, N., & Johari, P. K. (2015). A comparative study on feature extraction using texture and shape for content based image retrieval. *International Journal of Advanced Science and Technology*, 80(4), 41-52.
- Bai, X., Cao, Z., Zhao, L., Zhang, J., Lv, C., Li, C., & Xie, J. (2018). Rice heading stage automatic observation by multi-classifier cascade based rice spike detection method. *Agricultural and Forest Meteorology*, 259, 260-270.
- Barbedo, J. G. (2018). Factors influencing the use of deep learning for plant disease recognition. *Biosystems engineering*, 172, 84-91.
- Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems engineering*, 144, 52-60.
- Barburiceanu, S., Terebes, R., & Meza, S. (2020, May). Grape leaf disease classification using LBP-derived texture operators and colour. In *2020 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)* (pp. 1-6). IEEE.
- Basavaiah, J., & Arlene Anthony, A. (2020). Tomato leaf disease classification using multiple feature extraction techniques. *Wireless Personal Communications*, 115(1), 633-651.

Bhat, R. (Ed.). (2021). *Future Foods: Global Trends, Opportunities, and Sustainability Challenges*. Academic Press.

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

Bond, M. E., & Nickerson, D. (1942). Color-order systems, Munsell and Ostwald. *JOSA*, 32(12), 709-719.

Brahimi, M., Boukhalfa, K., & Moussaoui, A. (2017). Deep learning for tomato diseases: classification and symptoms visualization. *Applied Artificial Intelligence*, 31(4), 299-315.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Agrostat: estatísticas de comércio exterior do agronegócio brasileiro: indicadores gerais Agrostat. Disponível em: <http://indicadores.agricultura.gov.br/agrostat/index.htm>. Acesso em: 1 out. 2022.

Brynjolfsson, E., & McAfee, A. N. D. R. E. W. (2017). Artificial intelligence, for real. *Harvard business review*, 1, 1-31.

Burger, W., & Burge, M. J. (2022). *Digital Image Processing: An Algorithmic Introduction*. Springer Nature.

Buttan, Y., Chaudhary, A., & Saxena, K. (2021). An improved model for breast cancer classification using random forest with grid search method. In *Proceedings of Second International Conference on Smart Energy and Communication: ICSEC 2020* (pp. 407-415). Springer Singapore.

Cardoso, R. C., & Ferrando, A. (2021). A review of agent-based programming for multi-agent systems. *Computers*, 10(2), 16.

Carollo, E.M.; Filho, H.P. S. (2016). Manual básico de técnicas fitopatológicas. Cruz das Almas: Embrapa Mandioca e Fruticultura BA.

CEPEA/USP e CNA, (2020). Centro de Estudos Avançados em Economia Aplicada (Cepea), da Esalq/USP, em parceria com a Confederação da Agricultura e Pecuária do Brasil (CNA). Disponível em: <http://www.cepea.esalq.usp.br/br/pib-do-agronegocio-brasileiro.aspx>. Acessado no dia 04/11/2021: 12h:56min.

Chaki, J., & Dey, N. (2020). *Image color feature extraction techniques: fundamentals and applications*. Springer Nature.

Chen, F., Ye, Z., Wang, C., Yan, L., & Wang, R. (2018, September). A feature selection approach for network intrusion detection based on tree-seed algorithm and k-nearest neighbor. In *2018 IEEE 4th international symposium on wireless systems within the*

international conferences on intelligent data acquisition and advanced computing systems (IDAACS-SWS) (pp. 68-72). IEEE.

Chen, J., Chen, J., Zhang, D., Sun, Y., & Nanekaran, Y. A. (2020). Using deep transfer learning for image-based plant disease identification. *Computers and Electronics in Agriculture*, *173*, 105393.

Chen, J., Yin, H., & Zhang, D. (2020). A self-adaptive classification method for plant disease detection using GMDH-Logistic model. *Sustainable Computing: Informatics and Systems*, *28*, 100415.

Chen, R., Qi, H., Liang, Y., & Yang, M. (2022). Identification of plant leaf diseases by deep learning based on channel attention and channel pruning. *Frontiers in Plant Science*, *13*, 1023515.

Chen, S., Zhang, K., Zhao, Y., Sun, Y., Ban, W., Chen, Y., ... & Yang, T. (2021). An approach for rice bacterial leaf streak disease segmentation and disease severity estimation. *Agriculture*, *11*(5), 420.

Cheng, R. (2020). A survey: Comparison between Convolutional Neural Network and YOLO in image identification. In *Journal of Physics: Conference Series* (Vol. 1453, No. 1, p. 012139). IOP Publishing.

Chiu, W. T., Lin, C. H., Jhu, C. L., Lin, C., Chen, Y. C., Huang, M. J., & Liu, W. M. (2020, December). Semantic segmentation of lotus leaves in UAV aerial images via U-Net and deepLab-based networks. In *2020 International Computer Symposium (ICS)* (pp. 535-540). IEEE.

Chouhan, S. S., Singh, U. P., & Jain, S. (2020). Applications of computer vision in plant pathology: a survey. *Archives of computational methods in engineering*, *27*, 611-632.

Chouhan, S. S., Singh, U. P., & Jain, S. (2021). Automated plant leaf disease detection and classification using fuzzy based function network. *Wireless Personal Communications*, *121*, 1757-1779.

Cihan, P. (2020, October). Fuzzy rule-based system for predicting daily case in covid-19 outbreak. In *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)* (pp. 1-4). IEEE.

Cimpoi, M., Maji, S., Kokkinos, I., & Vedaldi, A. (2016). Deep filter banks for texture recognition, description, and segmentation. *IJCV*, *118*(1), 65–94.

Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., & Traore, D. (2019). Deep neural networks with transfer learning in millet crop images. *Computers in industry*, *108*, 115-120.

DeChant, C., Wiesner-Hanks, T., Chen, S., Stewart, E. L., Yosinski, J., Gore, M. A., ... & Lipson, H. (2017). Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. *Phytopathology*, *107*(11), 1426-1432.

Deleon, L., Brewer, M. J., Esquivel, I. L., & Halcomb, J. (2017). Use of a geographic information system to produce pest monitoring maps for south Texas cotton and sorghum land managers. *Crop Protection*, *101*, 50-57.

Devaraj, A., Rathan, K., Jaahnavi, S., & Indira, K. (2019, April). Identification of plant disease using image processing technique. In *2019 International Conference on Communication and Signal Processing (ICCSP)* (pp. 0749-0753). IEEE.

Dhaka, V. S., Meena, S. V., Rani, G., Sinwar, D., Ijaz, M. F., & Woźniak, M. (2021). A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. *Sensors*, *21*(14), 4749.

Dhar, S., Vishwakarma, A., Ghanti, D., & Jana, N. D. (2022, November). Ensemble Learning based Plant Leaf Disease Classification Considering Deep Convolutional Features from Pre-trained CNN. In *2022 IEEE 6th Conference on Information and Communication Technology (CICT)* (pp. 1-6). IEEE.

Dhaware, C. G., & Wanjale, K. H. (2017, January). A modern approach for plant leaf disease classification which depends on leaf image processing. In *2017 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-4). IEEE.

Dhinesh, E., & Jagan, A. (2019, December). Detection of Leaf Disease Using Principal Component Analysis and Linear Support Vector Machine. In *2019 11th International Conference on Advanced Computing (ICoAC)* (pp. 350-355). IEEE.

Dhingra, G., Kumar, V., & Joshi, H. D. (2019). A novel computer vision based neutrosophic approach for leaf disease identification and classification. *Measurement*, *135*, 782-794.

Doh, B., Zhang, D., Shen, Y., Hussain, F., Doh, R. F., & Ayepah, K. (2019, September). Automatic citrus fruit disease detection by phenotyping using machine learning. In *2019 25th International Conference on Automation and Computing (ICAC)* (pp. 1-5). IEEE.

Dong, Y., Xu, F., Liu, L., Du, X., Ye, H., Huang, W., & Zhu, Y. (2019, July). Monitoring and forecasting for disease and pest in crop based on WebGIS system. In *2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)* (pp. 1-5). IEEE.

Donkor, E., Jnr, E. M. O., Adu-Bredu, S., Andam-Akorful, S. A., Kwarteng, E. V. S., & Yevugah, L. L. (2022). Application of Parametric and Non Parametric Classifiers for

Assessing Land Use/Land Cover Categories in Cocoa Landscape of Juaboso and Bia West Districts of Ghana. *Journal of Geoscience and Environment Protection*, 10(11), 265-281.

Erhan, D., Szegedy, C., Toshev, A., & Anguelov, D. (2014). Scalable object detection using deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2147-2154).

Fan, X., Luo, P., Mu, Y., Zhou, R., Tjahjadi, T., & Ren, Y. (2022). Leaf image based plant disease identification using transfer learning and feature fusion. *Computers and Electronics in agriculture*, 196, 106892.

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and electronics in agriculture*, 145, 311-318.

Francis, J., & Anoop, B. K. (2016, March). Identification of leaf diseases in pepper plants using soft computing techniques. In *2016 conference on emerging devices and smart systems (ICEDSS)* (pp. 168-173). IEEE.

Fujita, E., Kawasaki, Y., Uga, H., Kagiwada, S., & Iyatomi, H. (2016, December). Basic investigation on a robust and practical plant diagnostic system. In *2016 15th IEEE international conference on machine learning and applications (ICMLA)* (pp. 989-992). IEEE.

Fulari, U. N., Shastri, R. K., & Fulari, A. N. (2020). Leaf disease detection using machine learning. *J. Seybold Rep*, 1533, 9211.

Gayathri Devi, T., & Neelamegam, P. (2019). Image processing based rice plant leaves diseases in Thanjavur, Tamilnadu. *Cluster Computing*, 22, 13415-13428.

Gharge, S., & Singh, P. (2016). Image processing for soybean disease classification and severity estimation. In *Emerging Research in Computing, Information, Communication and Applications: ERCICA 2015, Volume 2* (pp. 493-500). Springer India.

Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).

Golhani, K., Balasundram, S. K., Vadamalai, G., & Pradhan, B. (2018). A review of neural networks in plant disease detection using hyperspectral data. *Information Processing in Agriculture*, 5(3), 354-371.

Gonzales, R. C., Woods, R.E. (1992). Digital Image Processing, ISBN 0-201-50803-6.

Gonzales, R. C., Woods, R.E. (2010). Digital Image Processing, Third Edition, Pearson Education.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Govardhan, M., & Veena, M. B. (2019, October). Diagnosis of tomato plant diseases using random forest. In *2019 Global Conference for Advancement in Technology (GCAT)* (pp. 1-5). IEEE.

Grigoletti Júnior, A.; Auer, C. G. (2013). Doenças da erva-mate: identificação e controle. Circular técnica, Colombo, n. 25, 1996. Disponível em: http://ainfo.cnptia.embrapa.br/digital/bitstream/item/16966/1/circ_tec25.pdf. Acesso em: 4 abr. 2021.

Gunarathna, M. M., & Rathmayaa, R. M. K. T. (2020). Efficient deep learning models for tomato plant disease classification based on leaf image. In *International Conference on Advances in Computing and Technology*.

Gupta, D., & Singh, R. (2019). Disease Detection in Plant using Artificial Neural Network. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 7(VI), 2560-2565.

Hasan, R. I., Yusuf, S. M., Rahim, M. S. M., & Alzubaidi, L. (2022). Automated masks generation for coffee and apple leaf infected with single or multiple diseases-based color analysis approaches. *Informatics in Medicine Unlocked*, 28, 100837.

Hassan, S. M., & Maji, A. K. (2022). Plant disease identification using a novel convolutional neural network. *IEEE Access*, 10, 5390-5401.

Hassan, S. M., Maji, A. K., Jasiński, M., Leonowicz, Z., & Jasińska, E. (2021). Identification of plant-leaf diseases using CNN and transfer-learning approach. *Electronics*, 10(12), 1388.

Hatuwal, B. K., Shakya, A., & Joshi, B. (2020). Plant Leaf Disease Recognition Using Random Forest, KNN, SVM and CNN. *Polibits*, 62, 13-19.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).

Hlaing, C. S., & Zaw, S. M. M. (2018, June). Tomato plant diseases classification using statistical texture feature and color feature. In *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)* (pp. 439-444). IEEE.

Ho, T. K. (1995, August). Random decision forests. In Proceedings of 3rd international conference on document analysis and recognition (Vol. 1, pp. 278-282). IEEE.

Hossain, E., Hossain, M. F., & Rahaman, M. A. (2019, February). A color and texture based approach for the detection and classification of plant leaf disease using KNN classifier. In *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)* (pp. 1-6). IEEE.

Hu, G., Wu, H., Zhang, Y., & Wan, M. (2019). A low shot learning method for tea leaf's disease identification. *Computers and Electronics in Agriculture*, 163, 104852.

Huang, T., Yang, R., Huang, W., Huang, Y., & Qiao, X. (2018). Detecting sugarcane borer diseases using support vector machine. *Information processing in agriculture*, 5(1), 74-82.

Humeau-Heurtier, A. (2019). Texture feature extraction methods: A survey. *IEEE access*, 7, 8975-9000.

IBGE, (2021). Instituto Brasileiro de Geografia e Estatística. Estatística Paraná. 2021. Disponível em: <https://sidra.ibge.gov.br/tabela/5457#resultado>.

IBGE. Sidra: banco de tabelas estatísticas: índice nacional de preços ao consumidos amplo – setembro 2020. [Brasília, DF], 9 out. 2020. Disponível em: <https://sidra.ibge.gov.br/home/ipca/brasil>. Acesso em: 1 out. 2020.

Iqbal, M. A., & Talukder, K. H. (2020, August). Detection of potato disease using image segmentation and machine learning. In *2020 international conference on wireless communications signal processing and networking (WiSPNET)* (pp. 43-47). IEEE.

Jaiganesh, M., Sathyadevi, M., Chakravarthy, K. S., & Sarada, C. (2020). Identification of plant species using CNN-classifier. *Journal Of Critical Reviews*, 7(3), 923-931.

Jaisakthi, S. M., Mirunalini, P., & Thenmozhi, D. (2019, February). Grape leaf disease identification using machine learning techniques. In *2019 International conference on computational intelligence in data science (ICCIDS)* (pp. 1-6). IEEE.

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685-695.

Jayakumar, A. (2019). Detection and Classification of Leaf Diseases in Maize Plant using Machine Learning (Doctoral dissertation, Dublin, National College of Ireland).

Jia, Z., Shi, A., Xie, G., & Mu, S. (2022, April). Image segmentation of persimmon leaf diseases based on UNet. In *2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)* (pp. 2036-2039). IEEE.

Jiang, H., & Learned-Miller, E. (2017, May). Face detection with the faster R-CNN. In 2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017) (pp. 650-657). IEEE.

Jocher, G., Chaurasia, A., Qiu, J. 2023. Ultralytics YOLOv8. <https://github.com/ultralytics/ultralytics>.

Johannes, A., Picon, A., Alvarez-Gila, A., Echazarra, J., Rodriguez-Vaamonde, S., Navajas, A. D., & Ortiz-Barredo, A. (2017). Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. *Computers and electronics in agriculture*, 138, 200-209.

Joseph, A. A., Abdullahi, M., Junaidu, S. B., Ibrahim, H. H., & Chiroma, H. (2022). Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer). *Intelligent Systems with Applications*, 14, 200066.

Kamal, K. C., Yin, Z., Wu, M., & Wu, Z. (2019). Depthwise separable convolution architectures for plant disease classification. *Computers and electronics in agriculture*, 165, 104948.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, 147, 70-90.

Karthik, R., Hariharan, M., Anand, S., Mathikshara, P., Johnson, A., & Menaka, R. (2020). Attention embedded residual CNN for disease detection in tomato leaves. *Applied Soft Computing*, 86, 105933.

Kaur, N. (2021). Plant leaf disease detection using ensemble classification and feature extraction. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(11), 2339-2352.

Kaur, S., Pandey, S., & Goel, S. (2018). Semi-automatic leaf disease detection and classification system for soybean culture. *IET Image Processing*, 12(6), 1038-1048.

Kaur, V., & Oberoi, A. (2020). Novel Approach for Plant Disease Detection Based on Textural Feature Analysis. In *Data Management, Analytics and Innovation: Proceedings of ICDMAI 2019, Volume 1* (pp. 439-449). Springer Singapore.

Kawasaki, Y., Uga, H., Kagiwada, S., & Iyatomi, H. (2015). Basic study of automated diagnosis of viral plant diseases using convolutional neural networks. In *Advances in Visual Computing: 11th International Symposium, ISVC 2015, Las Vegas, NV, USA, December 14-16, 2015, Proceedings, Part II 11* (pp. 638-645). Springer International Publishing.

Khan, M. A., Akram, T., Sharif, M., Awais, M., Javed, K., Ali, H., & Saba, T. (2018). CCDF: Automatic system for segmentation and recognition of fruit crops diseases based on correlation coefficient and deep CNN features. *Computers and electronics in agriculture*, 155, 220-236.

Khatri, A., Agrawal, S., & Chatterjee, J. M. (2022). Wheat seed classification: utilizing ensemble machine learning approach. *Scientific Programming*, 2022.

Khilari, N., Hadawale, P., Shaikh, H., & Kolase, S. (2021). Analysis of Machine Learning Algorithm to predict Wine Quality.

Khitthuk, C., Srikaew, A., Attakitmongcol, K., & Kumsawat, P. (2018, March). Plant leaf disease diagnosis from color imagery using co-occurrence matrix and artificial intelligence system. In 2018 International Electrical Engineering Congress (iEECON) (pp. 1-4). IEEE.

Kimati, H., Amorim, L., Rezende, J. A. M., Bergamin Filho, A., & Camargo, L. E. A. (2005). Manual de Fitopatologia vol. 02–doenças das plantas cultivadas. São Paulo, SP: Editora Agronômica Ceres Ltda.

Kittler, H. et al. (2017). Diagnostic accuracy of dermoscopy. *The Lancet Oncology, Elsevier*, v. 3, n. 3, p. 159–165, dec 2017. ISSN 1470-2045. Disponível em: [http://dx.doi.org/10.1016/S1470-2045\(02\)00679-4](http://dx.doi.org/10.1016/S1470-2045(02)00679-4).

Koklu, M., & Ozkan, I. A. (2020). Multiclass classification of dry beans using computer vision and machine learning techniques. *Computers and Electronics in Agriculture*, 174, 105507.

Koschan, A., & Abidi, M. (2008). *Digital color image processing*. John Wiley & Sons.

Kumar, R. R., Cirrincione, G., Cirrincione, M., Andriollo, M., & Tortella, A. (2018, September). Accurate fault diagnosis and classification scheme based on non-parametric, statistical-frequency features and neural networks. In 2018 XIII International Conference on Electrical Machines (ICEM) (pp. 1747-1753). IEEE.

Kumar, R. R., Cirrincione, G., Cirrincione, M., Tortella, A., & Andriollo, M. (2020). Induction machine fault detection and classification using non-parametric, statistical-frequency features and shallow neural networks. *IEEE transactions on Energy Conversion*, 36(2), 1070-1080.

Kumar, R., Chug, A., Singh, A. P., & Singh, D. (2022). A Systematic analysis of machine learning and deep learning based approaches for plant leaf disease classification: a review. *Journal of Sensors*, 2022.

Kurmi, Y., Gangwar, S., Agrawal, D., Kumar, S., Saxena, D., Saxena, M., & Shrivastava, H. (2020, February). An algorithm for various crop diseases detection and

classification using leaves images. In *2nd International Conference on Data, Engineering and Applications (IDEA)* (pp. 1-5). IEEE.

Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, *14*(5), 778-782.

Lambora, A., Gupta, K., & Chopra, K. (2019, February). Genetic algorithm-A literature review. In *2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)* (pp. 380-384). IEEE.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.

Lecun, Y.; Bottou, L.; Orr, G. B.; Müller, K. R. (1998). Efficient backprop. In: *Neural Networks: Tricks of the Trade*. Berlin, Heidelberg: Springer Berlin Heidelberg, p. 9–50. Citado na página 17.

Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., ... & Wei, X. (2022). YOLOv6: A single-stage object detection framework for industrial applications. *arXiv preprint arXiv:2209.02976*.

Li, J., Chen, L., & Huang, W. (2018). Detection of early bruises on peaches (*Amygdalus persica* L.) using hyperspectral imaging coupled with improved watershed segmentation algorithm. *Postharvest Biology and Technology*, *135*, 104-113.

Li, W. (2021, March). Analysis of object detection performance based on Faster R-CNN. In *Journal of Physics: Conference Series* (Vol. 1827, No. 1, p. 012085). IOP Publishing.

Li, Y., Nie, J., & Chao, X. (2020). Do we really need deep CNN for plant diseases identification?. *Computers and Electronics in Agriculture*, *178*, 105803.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, *18*(8), 2674.

Liang, M., & Hu, X. (2015). Recurrent convolutional neural network for object recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3367-3375).

Lin, K., Gong, L., Huang, Y., Liu, C., & Pan, J. (2019). Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. *Frontiers in plant science*, *10*, 155.

Liu, F., Liu, Z., Liu, W., & Zhao, H. (2022, November). Combining the YOLOv5 and Grabcut Algorithms for Fashion Color Analysis of Clothing. In *2022 5th World*

Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM) (pp. 1126-1129). IEEE.

Liu, J., & Wang, X. (2021). Plant diseases and pests detection based on deep learning: a review. *Plant Methods*, 17, 1-18.

Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., & Pietikäinen, M. (2020). Deep learning for generic object detection: A survey. *International journal of computer vision*, 128, 261-318.

Liu, X., Xu, B., Gu, W., Yin, Y., & Wang, H. (2022). Plant leaf veins coupling feature representation and measurement method based on DeepLabV3+. *Frontiers in Plant Science*, 13, 1043884.

Lou, H., Duan, X., Guo, J., Liu, H., Gu, J., Bi, L., & Chen, H. (2023). DC-YOLOv8: Small-Size Object Detection Algorithm Based on Camera Sensor. *Electronics*, 12(10), 2323.

Lu, J., Ehsani, R., Shi, Y., Abdulridha, J., de Castro, A. I., & Xu, Y. (2017). Field detection of anthracnose crown rot in strawberry using spectroscopy technology. *computers and electronics in agriculture*, 135, 289-299.

Lu, J., Hu, J., Zhao, G., Mei, F., & Zhang, C. (2017). An in-field automatic wheat disease diagnosis system. *Computers and electronics in agriculture*, 142, 369-379.

Ma, J., Du, K., Zheng, F., Zhang, L., Gong, Z., & Sun, Z. (2018). A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. *Computers and electronics in agriculture*, 154, 18-24.

Madiwalar, S. C., & Wyawahare, M. V. (2017, February). Plant disease identification: A comparative study. In 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI) (pp. 13-18). IEEE.

Magsi, A., Mahar, J. A., Razzaq, M. A., & Gill, S. H. (2020, November). Date palm disease identification using features extraction and deep learning approach. In 2020 IEEE 23rd International Multitopic Conference (INMIC) (pp. 1-6). IEEE.

Marques, J. J., Burille, C., Schuster, C., Brandoli, M. A. A., Rocha, D., Bresolin, B., ... & Silveira, J. D. (2013). Erva-mate: guia para aplicação de boas práticas agrícolas. *Lajeado, RS: Emater/RS-Ascar*, 80.

Mathew, D., Kumar, C. S., & Cherian, K. A. (2021). Foliar fungal disease classification in banana plants using elliptical local binary pattern on multiresolution dual tree complex wavelet transform domain. *Information processing in Agriculture*, 8(4), 581-592.

Medeiros, A. D. D., Silva, L. J. D., Ribeiro, J. P. O., Ferreira, K. C., Rosas, J. T. F., Santos, A. A., & Silva, C. B. D. (2020). Machine learning for seed quality classification: An advanced approach using merger data from FT-NIR spectroscopy and X-ray imaging. *Sensors*, 20(15), 4319.

Medhi, S., Ahmed, C., & Gayan, R. (2016). A study on feature extraction techniques in image processing. *International Journal of Computer Sciences and Engineering*, 4(7), 89-93.

Minar, M. R., & Naher, J. (2018). Recent advances in deep learning: An overview. *arXiv preprint arXiv:1807.08169*.

Mohanapriya, K., & Balasubramani, M. (2019, October). Recognition of unhealthy plant leaves using Naive Bayes classifier. In *IOP Conference Series: Materials Science and Engineering* (Vol. 561, No. 1, p. 012094). IOP Publishing.

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, 7, 1419.

Nageswara Rao, P. P., & Lakshmikantha, B. P. (2020). Applications of geospatial technologies in plant health management. *Innovative Pest Management Approaches for the 21st Century: Harnessing Automated Unmanned Technologies*, 3-25.

Nair, L. and Saju, K. K. (2019). Classification of macronutrient deficiencies in maize plants using optimized multi class support vector machines, *Engineering in Agriculture*. 12 (1) 126-139.

Ochango, V. M., Wambugu, G. M., & Ndia, J. G. (2022). Feature extraction using histogram of oriented gradients for image classification in maize leaf diseases. *International Journal of Computer and Information Technology* (2279-0764), 11(2).

Padol, P. B., & Yadav, A. A. (2016, June). SVM classifier based grape leaf disease detection. In *2016 Conference on advances in signal processing (CASP)* (pp. 175-179). IEEE.

Parida, P., & Bhoi, N. (2017). 2-D Gabor filter based transition region extraction and morphological operation for image segmentation. *Computers & Electrical Engineering*, 62, 119-134.

Patidar, H.K., Vianny, M.M. (2020, July). Leaf Segmentation Challenge Using UNET. *International Journal of Creative Research Thoughts (IJCRT) 2020 IJCRT*, Volume 8, Issue 7, ISSN: 2320-2882.

Pedrini, H., & Schwartz, W. R. (2008). *Análise de imagens digitais: princípios, algoritmos e aplicações*. Cengage Learning.

- Phan, H., Ahmad, A., & Saraswat, D. (2022). Identification of Foliar Disease Regions on Corn Leaves Using SLIC Segmentation and Deep Learning Under Uniform Background and Field Conditions. *IEEE Access*, *10*, 111985-111995.
- Piassetta, R. D. R. L., Mikos, A. P., & Auer, C. G. (2021). DOENÇAS FÚNGICAS DA CULTURA DA ERVA-MATE (*Ilex paraguariensis*) NO BRASIL. *BIOFIX Scientific Journal*, *6*(2), 153-159.
- Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J., & Johannes, A. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture*, *161*, 280-290.
- Poletto, I., Muniz, M. F. B., Blume, E., & Ceconi, D. E. (2010). Efeito da aplicação de NPK e calcário sobre a podridão-de-raízes em mudas de erva-mate. *Floresta*, *40*(3).
- Pothen, M. E., & Pai, M. L. (2020, March). Detection of rice leaf diseases using image processing. In 2020 fourth international conference on computing methodologies and communication (ICCMC) (pp. 424-430). IEEE.
- Prajapati, H. B., Shah, J. P., & Dabhi, V. K. (2017). Detection and classification of rice plant diseases. *Intelligent Decision Technologies*, *11*(3), 357-373.
- Prasetyo, S. Y. J., Agus, Y. H., Dewi, C., Simanjuntak, B. H., & Hartomo, K. D. (2017, March). GEODATA: Information System Based on Geospatial for Early Warning Tracking and Analysis Agricultural Plant Diseases in Central Java. In *IOP Conference Series: Materials Science and Engineering* (Vol. 180, No. 1, p. 012070). IOP Publishing.
- Pratt, W. K. (2001). *Digital Image Processing*. PIKS inside, John Wiley & Sons. Inc Editions.
- Pujari, D., Yakkundimath, R., & Byadgi, A. S. (2016). SVM and ANN based classification of plant diseases using feature reduction technique. *IJIMAI*, *3*(7), 6-14.
- Qi, H., Liang, Y., Ding, Q., & Zou, J. (2021). Automatic identification of peanut-leaf diseases based on stack ensemble. *Applied Sciences*, *11*(4), 1950.
- Radhakrishnan, S. (2020). An improved machine learning algorithm for predicting blast disease in paddy crop. *Materials Today: Proceedings*, *33*, 682-686.
- Ramaniharan, A. K., Manoharan, S. C., & Swaminathan, R. (2016). Laplace Beltrami eigen value based classification of normal and Alzheimer MR images using parametric and non-parametric classifiers. *Expert Systems with Applications*, *59*, 208-216.

- Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., & Hughes, D. P. (2017). Deep learning for image-based cassava disease detection. *Frontiers in plant science*, 8, 1852.
- Ramesh, S., & Vydeki, D. (2020). Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm. *Information processing in agriculture*, 7(2), 249-260.
- Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Shashank, N., & Vinod, P. V. (2018, April). Plant disease detection using machine learning. In *2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C)* (pp. 41-45). IEEE.
- Rao, A., & Kulkarni, S. B. (2020). A Hybrid Approach for Plant Leaf Disease Detection and Classification Using Digital Image Processing Methods. *International Journal of Electrical Engineering Education*.
- Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7263-7271).
- Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
- Rehman, T. U., Mahmud, M. S., Chang, Y. K., Jin, J., & Shin, J. (2019). Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. *Computers and electronics in agriculture*, 156, 585-605.
- Reis, D., Kupec, J., Hong, J., & Daoudi, A. (2023). Real-Time Flying Object Detection with YOLOv8. *arXiv preprint arXiv:2305.09972*.
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.
- Roberts, D. P., Short, N. M., Sill, J., Lakshman, D. K., Hu, X., & Buser, M. (2021). Precision agriculture and geospatial techniques for sustainable disease control. *Indian Phytopathology*, 74, 287-305.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* (pp. 234-241). Springer International Publishing.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.

Rosenfeld, A., & Kak, A. (1982). Edge detection. In *Digital picture processing* (Vol. 2, pp. 84-112). Orlando, FL, USA: Academic press.

Rother, C., Kolmogorov, V., & Blake, A. (2004). " GrabCut" interactive foreground extraction using iterated graph cuts. *ACM transactions on graphics (TOG)*, 23(3), 309-314.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature machine intelligence*, 1(5), 206-215.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115, 211-252.

Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Saberi Anari, M. (2022). A hybrid model for leaf diseases classification based on the modified deep transfer learning and ensemble approach for agricultural aiot-based monitoring. *Computational Intelligence and Neuroscience*, 2022.

Sabtu, N. M., Idris, N. H., & Ishak, M. H. I. (2018, June). The role of geospatial in plant pests and diseases: an overview. In *IOP Conference Series: Earth and Environmental Science* (Vol. 169, No. 1, p. 012013). IOP Publishing.

Saha, S., & Ahsan, S. M. M. (2021, December). Rice Leaf Disease Recognition using Gray-Level Co-Occurrence Matrix and Statistical Features. In *2021 5th International Conference on Electrical Information and Communication Technology (EICT)* (pp. 1-5). IEEE.

Saputra, R. A., Wasiyanti, S., Saefudin, D. F., Supriyatna, A., & Wibowo, A. (2020, November). Rice leaf disease image classifications using KNN based on GLCM feature extraction. In *Journal of Physics: Conference Series* (Vol. 1641, No. 1, p. 012080). IOP Publishing.

Sehgal, A., & Mathur, S. (2019, June). Plant disease classification using soft computing supervised machine learning. In *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 75-80). IEEE.

Selvaraj, M. G., Vergara, A., Ruiz, H., Safari, N., Elayabalan, S., Ocimati, W., & Blomme, G. (2019). AI-powered banana diseases and pest detection. *Plant methods*, 15, 1-11.

Shah, N., & Jain, S. (2019, February). Detection of disease in cotton leaf using artificial neural network. In *2019 Amity International Conference on Artificial Intelligence (AICAI)* (pp. 473-476). IEEE.

Sharif, M., Khan, M. A., Iqbal, Z., Azam, M. F., Lali, M. I. U., & Javed, M. Y. (2018). Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. *Computers and electronics in agriculture*, *150*, 220-234.

Sharma, S., & Gupta, A. (2020). A review for the automatic methods of plant's leaf image segmentation. *International Journal of Intelligence and Sustainable Computing*, *1*(1), 101-114.

Shetty, B., Fernandes, R., Rodrigues, A. P., Chengoden, R., Bhattacharya, S., & Lakshmana, K. (2022). Skin lesion classification of dermoscopic images using machine learning and convolutional neural network. *Scientific Reports*, *12*(1), 18134.

Shi, T., Liu, Y., Zheng, X., Hu, K., Huang, H., Liu, H., & Huang, H. (2023). Recent advances in plant disease severity assessment using convolutional neural networks. *Scientific Reports*, *13*(1), 2336.

Shingade, S. D., Mudhalwadkar, R. P., & Masal, K. M. (2022, October). Random Forest Machine Learning Classifier for Seed Recommendation. In *2022 International Conference on Edge Computing and Applications (ICECAA)* (pp. 1385-1390). IEEE.

Simón, D., Borsani, O., & Filippi, C. V. (2022). RFPDR: a random forest approach for plant disease resistance protein prediction. *PeerJ*, *10*, e11683.

Singh, A. K., Sreenivasu, S. V. N., Mahalaxmi, U. S. B. K., Sharma, H., Patil, D. D., & Asenso, E. (2022). Hybrid feature-based disease detection in plant leaf using convolutional neural network, bayesian optimized SVM, and random forest classifier. *Journal of Food Quality*, *2022*, 1-16.

Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, *97*, 105524.

Singh, M., Vermaa, A., & Kumar, V. (2023). Geospatial technologies for the management of pest and disease in crops. In *Precision Agriculture* (pp. 37-54). Academic Press.

Skala, V. (2023). Multispectral Image Generation from RGB Based on WSL Color Representation: Wavelength, Saturation, and Lightness. *Computers*, *12*(9), 182.

Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural networks based recognition of plant diseases by leaf image classification. *Computational intelligence and neuroscience*, *2016*.

Sobiyaa, P., Jayareka, K. S., Maheshkumar, K., Naveena, S., & Rao, K. S. (2022). Paddy disease classification using machine learning technique. *Materials Today: Proceedings*, 64, 883-887.

Stipaničev, D. (1994). *Introduction to Digital Image Processing and Analysis*. Chronolab Vision, Chronolab A.G. Zug Switzerland. R.Boskovica bb, 21000 SPLIT, Croatia , pp.1-78.

Storey, G., Meng, Q., & Li, B. (2022). Leaf disease segmentation and detection in apple orchards for precise smart spraying in sustainable agriculture. *Sustainability*, 14(3), 1458.

Subramani, B., Arumugam, S., & Ragupathy, B. (2019). Computer aided classification and detection of leaf disease using ANN. *Int J Innov Technol Explor Eng*, 8.

Sujatha, R., Chatterjee, J. M., Jhanjhi, N. Z., & Brohi, S. N. (2021). Performance of deep learning vs machine learning in plant leaf disease detection. *Microprocessors and Microsystems*, 80, 103615.

Suresha, M., Shreekanth, K. N., & Thirumalesh, B. V. (2017, April). Recognition of diseases in paddy leaves using knn classifier. In *2017 2nd International Conference for Convergence in Technology (I2CT)* (pp. 663-666). IEEE.

Szegedy, C., Reed, S., Erhan, D., Anguelov, D., & Ioffe, S. (2014). Scalable, high-quality object detection. *arXiv preprint arXiv:1412.1441*.

Tharwat, A. (2020). Classification assessment methods. *Applied computing and informatics*, 17(1), 168-192.

Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*, 161, 272-279.

Türkoğlu, M., & Hanbay, D. (2019). Plant disease and pest detection using deep learning-based features. *Turkish Journal of Electrical Engineering and Computer Sciences*, 27(3), 1636-1651.

UNICEF. (2021). The state of food security and nutrition in the world 2021.

UNITED STATES. Department of Agriculture. Foreign Agricultural Service. Custom query. Disponível em: <https://apps.fas.usda.gov/psdonline/app/index.html#/app/advQuery>. Acesso em: 1 out. 2020.

Ünver, H. M., & Ayan, E. (2019). Skin lesion segmentation in dermoscopic images with combination of YOLO and grabcut algorithm. *Diagnostics*, 9(3), 72.

Vamsidhar, E., Rani, P. J., & Babu, K. R. (2019). Plant disease identification and classification using image processing. *Int. J. Eng. Adv. Technol*, 8(3), 442-446.

VII Plano Diretor da Embrapa : 2020–2030 / Embrapa. – Brasília, DF : Embrapa, 2020. 31 p. : il. color. ; ? cm x ? cm.

Vishnoi, V. K., Kumar, K., & Kumar, B. (2022). A comprehensive study of feature extraction techniques for plant leaf disease detection. *Multimedia Tools and Applications*, 1-53.

Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7464-7475).

Wang, C., Du, P., Wu, H., Li, J., Zhao, C., & Zhu, H. (2021). A cucumber leaf disease severity classification method based on the fusion of DeepLabV3+ and U-Net. *Computers and Electronics in Agriculture*, 189, 106373.

Wu, H., Liu, Y., Xu, X., & Gao, Y. (2022). Object Detection Based on the GrabCut Method for Automatic Mask Generation. *Micromachines*, 13(12), 2095.

Xian, T. S., & Ngadiran, R. (2021, July). Plant diseases classification using machine learning. In *Journal of Physics: Conference Series* (Vol. 1962, No. 1, p. 012024). IOP Publishing.

Xu, S. (2018). Bayesian Naïve Bayes classifiers to text classification. *Journal of Information Science*, 44(1), 48-59.

Yalcin, H. (2017, August). Plant phenology recognition using deep learning: Deep-Pheno. In *2017 6th International Conference on Agro-Geoinformatics* (pp. 1-5). IEEE.

Yang, K., Zhong, W., & Li, F. (2020). Leaf segmentation and classification with a complicated background using deep learning. *Agronomy*, 10(11), 1721.

Yang, X., & Guo, T. (2017). Machine learning in plant disease research. *March*, 31, 1.

Yogeshwari, M., & Thailambal, G. (2021). Automatic feature extraction and detection of plant leaf disease using GLCM features and convolutional neural networks. *Materials Today: Proceedings*.

Yogeswararao, G., Malmathanraj, R., & Palanisamy, P. (2022). Fractional weighted nuclear norm based two dimensional linear discriminant features for cucumber leaf disease recognition. *Multimedia Tools and Applications*, 81(27), 38735-38755.

Young, I. T., Gerbrands, J. J., & Van Vliet, L. J. (1998). *Fundamentals of image processing* (Vol. 841). Delft: Delft University of Technology.

- Zapolska, A., Kalaitzidis, C., Markakis, E., Ligoxigakis, E., & Koubouris, G. (2020). Linear Discriminant Analysis of spectral measurements for discrimination between healthy and diseased trees of *Olea europaea* L. artificially infected by *Fomitiporia mediterranea*. *International Journal of Remote Sensing*, 41(14), 5388-5398.
- Zhang, J. C., Pu, R. L., Wang, J. H., Huang, W. J., Yuan, L., & Luo, J. H. (2012). Detecting powdery mildew of winter wheat using leaf level hyperspectral measurements. *Computers and Electronics in Agriculture*, 85, 13-23.
- Zhang, S., Wang, H., Huang, W., & You, Z. (2018). Plant diseased leaf segmentation and recognition by fusion of superpixel, K-means and PHOG. *Optik*, 157, 866-872.
- Zhang, S., Wu, X., You, Z., & Zhang, L. (2017). Leaf image based cucumber disease recognition using sparse representation classification. *Computers and electronics in agriculture*, 134, 135-141.
- Zhang, S., Zhang, S., Zhang, C., Wang, X., & Shi, Y. (2019). Cucumber leaf disease identification with global pooling dilated convolutional neural network. *Computers and Electronics in Agriculture*, 162, 422-430.
- Zhang, T., Fan, S., Xiang, Y., Zhang, S., Wang, J., & Sun, Q. (2020). Non-destructive analysis of germination percentage, germination energy and simple vigour index on wheat seeds during storage by Vis/NIR and SWIR hyperspectral imaging. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 239, 118488.
- Zhang, Y., & Ling, C. (2018). A strategy to apply machine learning to small datasets in materials science. *Npj Computational Materials*, 4(1), 25.
- Zhang, Z., Han, D., & Yang, Y. (2015, October). Image segmentation based on evidential Markov random field model. In *2015 International Conference on Control, Automation and Information Sciences (ICCAIS)* (pp. 239-244). IEEE.
- Zheng, Q., Huang, W., Cui, X., Dong, Y., Shi, Y., Ma, H., & Liu, L. (2018). Identification of wheat yellow rust using optimal three-band spectral indices in different growth stages. *Sensors*, 19(1), 35.
- Zheng, Q., Li, H., Fan, B., Wu, S., Xu, J., & Cao, Z. (2017). Modified localized multiplicative graph cuts based active contour model for object segmentation based on dynamic narrow band scheme. *Biomedical Signal Processing and Control*, 33, 119-131.
- Zivkovic, M., K, V., Bacanin, N., Djordjevic, A., Antonijevic, M., Strumberger, I., & Rashid, T. A. (2021, March). Hybrid genetic algorithm and machine learning method for covid-19 cases prediction. In *Proceedings of international conference on*

sustainable expert systems: ICSES 2020 (pp. 169-184). Singapore: Springer Singapore.

APÊNDICES – MATRIZES DE CONFUSÃO DOS EXPERIMENTOS

EXPERIMENTO A

TABELA 16 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (A) - CLASSIFICADOR RF

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	99	40	42	20	49,3
Pinta-preta	30	124	30	15	62,3
Fumagina	39	20	81	30	47,6
Requeima	29	30	34	86	48,0
Precisão (%)	50,3	57,9	43,3	56,9	-

Fonte: Autor (2023).

TABELA 17 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (A) - CLASSIFICADOR SVM

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	61	42	53	13	36,1
Pinta-preta	17	138	38	15	66,3
Fumagina	26	61	93	19	46,7
Requeima	28	50	42	58	32,6
Precisão (%)	46,2	47,4	41,2	55,2	-

Fonte: Autor (2023).

TABELA 18 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (A) - CLASSIFICADOR KNN

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	101	40	30	26	51,3
Pinta-preta	40	90	47	18	46,2
Fumagina	52	30	66	19	39,5
Requeima	48	42	48	49	26,2
Precisão (%)	41,9	44,5	34,6	43,8	-

Fonte: Autor (2023).

EXPERIMENTO B

TABELA 19 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (B) - CLASSIFICADOR RF

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	158	14	36	8	73,1
Pinta-preta	24	129	14	28	66,2
Fumagina	52	32	95	30	45,5
Requeima	14	28	10	130	71,4
Precisão (%)	63,7	63,5	61,3	66,3	-

Fonte: Autor (2023).

TABELA 20 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (B) - CLASSIFICADOR SVM

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	182	20	10	7	83,1
Pinta-preta	40	100	30	23	51,8
Fumagina	36	30	63	29	39,8
Requeima	20	40	9	120	63,5
Precisão (%)	65,5	52,6	56,3	67,0	-

Fonte: Autor (2023).

TABELA 21 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (B) - CLASSIFICADOR KNN

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	140	34	30	9	65,7
Pinta-preta	37	110	20	18	59,5
Fumagina	50	54	88	19	41,7
Requeima	20	57	31	96	47,1
Precisão (%)	56,7	43,1	52,1	67,6	-

Fonte: Autor (2023).

EXPERIMENTO C

TABELA 22 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (C) - CLASSIFICADOR RF

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	182	5	13	3	89,7
Pinta-preta	10	171	10	9	85,5
Fumagina	8	8	169	8	87,6
Requeima	6	12	12	160	84,2
Precisão (%)	88,3	87,2	82,8	88,8	-

Fonte: Autor (2023).

TABELA 23 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (C) - CLASSIFICADOR SVM

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	182	5	13	3	89,7
Pinta-preta	10	171	7	9	86,8
Fumagina	8	8	169	8	87,6
Requeima	6	12	10	160	85,1
Precisão (%)	88,4	87,2	84,9	88,8	-

Fonte: Autor (2023).

TABELA 24 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (C) - CLASSIFICADOR KNN

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	174	10	10	1	89,2
Pinta-preta	24	132	12	23	69,1
Fumagina	15	11	115	12	75,2
Requeima	15	21	13	111	69,4
Precisão (%)	76,3	75,8	76,7	75,5	-

Fonte: Autor (2023).

EXPERIMENTO D

TABELA 25 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (D) - CLASSIFICADOR RF

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	164	17	21	2	80,4
Pinta-preta	27	134	15	27	66,0
Fumagina	51	30	67	33	37,0
Requeima	12	26	7	135	75,0
Precisão (%)	64,6	64,7	60,9	68,6	

Fonte: Autor (2023).

TABELA 26 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (D) - CLASSIFICADOR SVM

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	163	8	17	2	85,8
Pinta-preta	42	90	20	17	53,3
Fumagina	63	21	66	28	37,1
Requeima	16	25	18	80	57,6
Precisão (%)	57,4	62,5	54,5	62,9	-

Fonte: Autor (2023).

TABELA 27 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (D) - CLASSIFICADOR KNN

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	139	22	27	9	70,6
Pinta-preta	50	115	20	19	56,4
Fumagina	45	48	71	25	37,6
Requeima	24	34	23	88	52,1
Precisão (%)	53,8	52,5	50,4	62,4	-

Fonte: Autor (2023).

EXPERIMENTO E

TABELA 28 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (E) - CLASSIFICADOR RF

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	186	6	8	4	91,2
Pinta-preta	10	168	10	11	84,0
Fumagina	10	10	148	7	84,6
Requeima	7	9	8	167	87,4
Precisão (%)	87,3	87,0	85,1	88,4	-

Fonte: Autor (2023).

TABELA 29 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (E) - CLASSIFICADOR SVM

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	181	10	12	3	87,8
Pinta-preta	24	127	12	20	69,0
Fumagina	17	19	118	10	71,9
Requeima	12	17	11	116	74,4
Precisão (%)	77,4	73,4	77,1	77,8	-

Fonte: Autor (2023).

TABELA 30 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (E) - CLASSIFICADOR KNN

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	149	30	14	5	75,3
Pinta-preta	26	142	12	28	68,0
Fumagina	24	37	94	19	54,0
Requeima	30	32	22	89	51,4
Precisão (%)	65,1	58,9	66,2	63,1	-

Fonte: Autor (2023).

EXPERIMENTO F

TABELA 31 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (F) - CLASSIFICADOR RF

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	191	7	13	2	89,7
Pinta-preta	8	160	16	11	82,0
Fumagina	12	9	152	7	84,4
Requeima	4	6	9	159	89,3
Precisão (%)	88,8	87,9	80,0	88,8	-

Fonte: Autor (2023).

TABELA 32 -MATRIZ DE CONFUSÃO DE EXPERIMENTO (F) - CLASSIFICADOR SVM

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	171	10	15	3	85,9
Pinta-preta	32	120	19	10	66,0
Fumagina	22	19	109	11	67,7
Requeima	15	25	18	110	65,5
Precisão (%)	71,3	68,9	68,0	82,1	-

Fonte: Autor (2023).

TABELA 33 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (F) - CLASSIFICADOR KNN

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	140	16	22	6	76,1
Pinta-preta	30	119	13	22	65,0
Fumagina	26	25	88	9	59,5
Requeima	23	41	16	92	53,5
Precisão (%)	63,9	59,2	63,0	71,3	-

Fonte: Autor (2023).

EXPERIMENTO G

TABELA 34 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (G) - CLASSIFICADOR RF

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	201	10	10	10	87,0
Pinta-preta	14	200	13	9	85,0
Fumagina	11	13	196	12	84,5
Requeima	7	10	14	199	86,5
Precisão (%)	86,3	85,8	84,0	86,5	-

Fonte: Autor (2023).

TABELA 35 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (G) - CLASSIFICADOR SVM

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	168	19	25	19	72,7
Pinta-preta	25	166	26	14	72,0
Fumagina	20	25	158	20	70,9
Requeima	20	23	24	166	71,2
Precisão (%)	72,1	72,2	68,0	75,8	-

Fonte: Autor (2023).

TABELA 36 - MATRIZ DE CONFUSÃO DE EXPERIMENTO (G) - CLASSIFICADOR KNN

	Saudáveis	Pinta-preta	Fumagina	Requeima	Recall (%)
Saudáveis	167	20	36	18	69,3
Pinta-preta	17	134	20	27	68,0
Fumagina	26	33	147	34	61,3
Requeima	23	46	30	160	61,8
Precisão (%)	71,7	57,5	63,0	66,9	-

Fonte: Autor (2023).