

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA - CCET  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA - DEE

CÁSSIO CIONI CARMO

**DETECÇÃO E IDENTIFICAÇÃO DE DOENÇAS EM  
FOLHAS UTILIZANDO REDES NEURAIS**

SÃO CARLOS -SP  
[2023]

CÁSSIO CIONI CARMO

**DETECÇÃO E IDENTIFICAÇÃO DE DOENÇAS EM FOLHAS UTILIZANDO REDES  
NEURAIS**

Trabalho de conclusão de curso apresentado ao Departamento de Engenharia Elétrica da Universidade Federal de São Carlos, para obtenção do título de bacharel em Engenharia Elétrica.

Orientador: Prof.Dr. Celso Ap.de França

São Carlos-SP  
[2023]

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia - CCET

Programa de Graduação em Engenharia Elétrica

**Comissão avaliadora.**

Membros da comissão examinadora que avaliou e aprovou a Defesa do Trabalho de Conclusão de Curso do candidato Cássio Cioni Carmo, realizada em 06/09/2023:

Prof. Dr. Celso Aparecido de França

Instituição: Universidade Federal de São Carlos

Prof. Dr. Luís Alberto Mijam Barêa

Instituição: Universidade Federal de São Carlos

Prof. Dr. Robson Barcellos

Instituição: Universidade Federal de São Carlos

## **AGRADECIMENTO**

Agradeço de coração a todos que me ajudaram nesta jornada acadêmica. Aos meus pais, obrigado por estarem sempre ao meu lado com amor e apoio. Talitha, minha noiva, sua compreensão e incentivo foram essenciais.

Aos professores do Departamento de Elétrica, especialmente aos Professores Celso e Robson, agradeço imensamente pela orientação valiosa que proporcionaram. Suas contribuições foram essenciais para o desenvolvimento deste trabalho. Obrigado pelo apoio e conhecimentos compartilhados.

Aos amigos e familiares, obrigado pelo apoio.

## RESUMO

Neste estudo, foi explorada uma abordagem com redes neurais para a detecção e classificação de doenças em folhas de plantas de soja com base em imagens. A pesquisa teve como objetivo principal desenvolver um modelo capaz de analisar imagens de folhas de soja, identificando diferentes classes de doenças e determinando a saúde geral das plantas. A importância desse trabalho reside na aplicação da agricultura de precisão, visando o monitoramento e cuidado mais eficaz das culturas. A metodologia empregada incluiu a extração de características das texturas das folhas, utilizando técnicas como *Histogram of Oriented Gradients* (HOG) para a identificação de padrões texturais e bordas, bem como a obtenção de histogramas de cores nas componentes HSV (*Hue*, *Saturation* e *Brightness*) e RGB (*Red*, *Green*, and *Blue*) das imagens. Os resultados obtidos demonstraram um desempenho promissor do modelo proposto na classificação das diferentes classes de doenças e na determinação da saúde das plantas de soja. Uma acurácia de 87.54% na classificação foi alcançada, indicando uma sólida taxa de classificação correta, mesmo diante da complexidade das classes e da variabilidade nas imagens. Em síntese, esta pesquisa oferece valiosas perspectivas para a detecção de doenças em plantas de soja por meio das redes neurais. Os resultados reforçam a viabilidade e relevância dessa abordagem para a agricultura de precisão e o monitoramento da saúde das culturas agrícolas. A análise da matriz de confusão revela informações valiosas sobre áreas que podem ser aprimoradas, especialmente nas classes com menor representatividade, enfatizando a importância de futuras otimizações. Isso abrange desde o aumento da coleta de dados e a expansão do conjunto de imagens de treinamento até ajustes mais detalhados nos hiperparâmetros das redes neurais. No futuro, a exploração de redes neurais convolucionais e o enriquecimento da base de dados têm o potencial de contribuir ainda mais para a precisão e eficácia do modelo, impulsionando avanços significativos na detecção e classificação de doenças em plantas.

**Palavras-chave:** Doenças de folhas. Redes Neurais. Visão Computacional.

## ABSTRACT

In this study, an approach using neural networks for the detection and classification of diseases in soybean leaves based on images was explored. The main objective of the research was to develop a model capable of analyzing soybean leaf images, identifying different disease classes, and determining the overall health of the plants. The significance of this work lies in its application to precision agriculture, aiming for more effective monitoring and care of crops. The methodology employed included the extraction of features from leaf textures, using techniques such as Histogram of Oriented Gradients (HOG) for identifying textural patterns and edges, as well as obtaining color histograms in the HSV (Hue, Saturation, and Brightness) and RGB (Red, Green, and Blue) components of the images. The results obtained demonstrated promising performance of the proposed model in classifying different disease classes and determining the health of soybean plants. An accuracy rate of 87.54% in classification was achieved, indicating a solid rate of correct classification, even in the face of the complexity of the classes and variability in the images. In summary, this research offers valuable insights for disease detection in soybean plants through neural networks. The results underscore the feasibility and relevance of this approach for precision agriculture and crop health monitoring. The analysis of the confusion matrix reveals valuable information about areas that can be improved, especially in classes with lower representation, emphasizing the importance of future optimizations. This encompasses increasing data collection and expanding the training image dataset, as well as making more detailed adjustments to the hyperparameters of the neural networks. In the future, the exploration of convolutional neural networks and the enrichment of the database have the potential to further contribute to the accuracy and effectiveness of the model, driving significant advancements in disease detection and classification in plants.

**Keyword:** Leaf Diseases. Neural Networks. Computer Vision.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Representação esquemática das funções básicas da planta e sintomas causados por alguns tipos de doenças	17
Figura 2 – Modelo matemático de um neurônio	21
Figura 3 – Sistema de Cores RGB	25
Figura 4 – Sistema de Cores HSV	26
Figura 5 – Diagrama de blocos do Sistema proposto	35
Figura 6 – Exemplar da doença Ferrugem	36
Figura 7 – Exemplar da doença Míldio	37
Figura 8 – Exemplar da doença Oídio	37
Figura 9 – Exemplar da doença Carijó	38
Figura 10 – Exemplar da doença Crestamento	38
Figura 11 – Exemplar da folha saudável	39
Figura 12 – Exemplificação do Blocos e Células	42
Figura 13 – Imagem de uma folha com doença Carijó com duas regiões destacadas	43
Figura 14 – HOG obtido para a região 1 da Figura 12	44
Figura 15 – HOG obtido para a região 2 da Figura 12	44
Figura 16 – Histogramas RGB obtido para a região 1 da Figura 12	45
Figura 17 – Histogramas RGB obtido para a região 2 da Figura 12	46
Figura 18 – Histograma HSV obtido para a região 1 da Figura 12	46
Figura 19 – Histograma HSV obtido para a região 2 da Figura 12	47
Figura 20 – Representação esquemática da rede neural escolhida	49
Figura 21 – Exemplar de folha com a doença Ferrugem	50
Figura 22 – Folha com Ferrugem no sistema de cores HSV	51
Figura 23 – Folha com Ferrugem na cor S	52
Figura 24 – Máscara obtida após filtragem	52
Figura 25 – Imagem obtida após aplicação da máscara	53
Figura 26 – Imagem original com os quadrados que foram separados para análise	54
Figura 27 – Resultados das predições de cada quadrado	55

## LISTA DE TABELAS

Tabela 1 – Número de texturas utilizadas por classe para treino	40
Tabela 2 – Número de texturas utilizadas por classe para teste	56
Tabela 3 – Acurácia de diferentes arquiteturas da camada oculta.	58
Tabela 4 – Matriz confusão obtida.	60
Tabela 5 – Métricas por classe	62
Tabela 6 – Comparativo da precisão	63



## LISTA DE ABREVIATURAS E SIGLAS

CNN - Rede Neural Convolucional (*Convolutional Neural Network*)

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária

HOG - Histograma de Gradientes Orientados (*Histogram of Oriented Gradient*)

HSV - Matiz, Saturação e Valor (*Hue Saturation Value*)

RGB - Vermelho, Verde e Azul (*Red Green Blue*)

RNA - Rede Neural Artificial

## SUMÁRIO

---

<b>1 INTRODUÇÃO</b>	<b>12</b>
1.1 CONTEXTUALIZAÇÃO	12
1.2 OBJETIVO	13
<b>1.2.1 Objetivo Geral</b>	<b>13</b>
<b>1.2.2 Objetivos Específicos</b>	<b>13</b>
1.3 JUSTIFICATIVA	14
1.4 MATERIAIS E MÉTODO	14
1.5 ESTRUTURA DO TRABALHO	15
<b>2 FUNDAMENTAÇÃO TEÓRICA E REVISÃO DE LITERATURA</b>	<b>17</b>
2.1 DOENÇA DE PLANTAS	17
2.2 DETECÇÃO DE DOENÇAS EM FOLHAS	19
2.3 REDES NEURAIS ARTIFICIAIS	21
2.4 VISÃO COMPUTACIONAL	23
2.5 IMAGEM DIGITAL	24
<b>2.5.1 Sistemas de Cores</b>	<b>26</b>
2.5.1.1 Sistema de Cores RGB	26
2.5.1.2 Sistema de Cores HSV	28
<b>2.5.2 HOG</b>	<b>29</b>
<b>2.5.3 Segmentação de Imagens</b>	<b>30</b>
<b>2.5.4 Processamento de Imagens</b>	<b>32</b>
2.6 TRABALHOS CORRELATOS	33
<b>3 DESENVOLVIMENTO DO TRABALHO</b>	<b>38</b>
3.1 BANCO DE IMAGENS	39
3.2 SELEÇÃO DE IMAGENS	42
3.3 EXTRAÇÃO DAS CARACTERÍSTICAS	43
3.4 TREINAMENTO DA REDE NEURAL	51
3.5 SEGMENTAÇÃO E PREDIÇÃO	53
<b>4 RESULTADOS</b>	<b>60</b>
4.1 ACURÁCIA	60
4.2 PRECISÃO	63
4.3 RECALL	63
4.3 F1-SCORE	63
4.4 MATRIZ DE CONFUSÃO	64
4.5 COMPARATIVO COM OUTROS TRABALHOS	67
<b>5 CONSIDERAÇÕES FINAIS</b>	<b>69</b>

5.1 CONCLUSÕES	69
5.2 TRABALHOS FUTUROS	69
<b>REFERÊNCIAS</b>	<b>71</b>

# 1 INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO

O processamento de imagem tem se destacado como uma ferramenta poderosa na detecção e identificação de doenças em diversas áreas, incluindo a agricultura. No contexto brasileiro, a importância da agricultura é inegável, sendo o país reconhecido mundialmente como um dos maiores produtores e exportadores de commodities agrícolas. Nesse sentido, a utilização de tecnologias avançadas, como o processamento de imagem, torna-se crucial para garantir a eficiência e a sustentabilidade do setor agrícola brasileiro.

A agricultura brasileira enfrenta uma série de desafios, incluindo o surgimento e a disseminação de doenças que afetam a saúde das plantas e comprometem a produtividade das lavouras. A detecção precoce e o diagnóstico preciso dessas doenças são fundamentais para a adoção de medidas de controle e mitigação, evitando perdas significativas na produção e reduzindo o impacto ambiental.

Nesse contexto, o processamento de imagem aplicado à análise de folhas surge como uma ferramenta promissora para auxiliar os agricultores e especialistas na identificação rápida e confiável de doenças vegetais. Ao automatizar o processo de detecção e identificação de doenças em folhas, o processamento de imagem pode contribuir para uma tomada de decisão mais eficiente e assertiva no manejo das plantações, permitindo a aplicação direcionada de defensivos agrícolas e a implementação de práticas de controle mais sustentáveis.

Além disso, a aplicação do processamento de imagem na detecção de doenças em folhas pode contribuir para a melhoria da qualidade e da segurança dos alimentos produzidos no Brasil. A identificação precoce de doenças permite um controle mais efetivo, reduzindo a necessidade de uso excessivo de agrotóxicos e garantindo a produção de alimentos mais saudáveis para a população.

## 1.2 OBJETIVO

### 1.2.1 Objetivo Geral

Desenvolvimento de um Sistema para detecção e classificação de doenças em folhas da lavoura brasileira por meio do processamento de imagens utilizando Redes Neurais.

### 1.2.2 Objetivos Específicos

Com base no objetivo geral do trabalho, foram estabelecidos os seguintes objetivos específicos:

- Pesquisa sobre as características das técnicas aplicáveis para a detecção de doenças em folhas, com a seleção daquela que melhor se adequa ao contexto do estudo;
- Identificação e análise das principais técnicas e algoritmos utilizados para detecção de doenças em folhas;
- Coleta de um conjunto de imagens de folhas saudáveis e infectadas com diferentes doenças;
- Pré-processamento das imagens para extração das características relevantes para a detecção de doenças;
- Implementação e treinamento da rede com as imagens pré-processadas.
- Avaliação e análise dos resultados obtidos;
- Otimizações para aprimoração do desempenho;
- Publicação dos resultados finais.

## 1.3 JUSTIFICATIVA

A justificativa para a realização desta monografia se baseia na relevância da detecção eficiente e precoce de doenças em folhas na lavoura. As doenças fitopatológicas representam um desafio significativo para a agricultura, pois podem resultar em perdas consideráveis de produtividade e qualidade dos cultivos. A identificação oportuna dessas doenças possibilita a implementação de estratégias de controle adequadas, minimizando os impactos negativos na produção agrícola

(GARRIDO et al., 2015).

A detecção manual de doenças em folhas é um processo lento e suscetível a erros. Nesse contexto, o uso de tecnologias avançadas, como o processamento de imagem e a inteligência artificial, tem demonstrado grande potencial para automatizar essa tarefa. A aplicação de algoritmos de detecção e classificação de doenças em folhas, com base em redes neurais, por exemplo, permite agilizar a identificação e proporcionar uma análise mais precisa das enfermidades (CHAUDHARY et al., 2012).

Além disso, a agricultura brasileira desempenha um papel fundamental na economia do país, sendo responsável por uma parcela expressiva da produção de alimentos. O aprimoramento das técnicas de detecção de doenças em folhas nas lavouras brasileiras é crucial para aumentar a eficiência e a competitividade do setor agrícola nacional (Confederação da Agricultura e Pecuária do Brasil, 2021).

#### 1.4 MATERIAIS E MÉTODO

A metodologia utilizada para realização do presente trabalho foi:

- Levantamento bibliográfico de metodologias de detecção de doenças e folhas por processamento de imagem, com consultas em artigos, livros, dissertações, trabalhos de conclusão de curso e monografias;
- Desenvolvimento do sistema de classificação utilizando redes neurais;
- Testes de validação;

Foram utilizados os seguintes materiais no trabalho:

- Software Visual Studio Code (ambiente de desenvolvimento de Python);
- Repositório Digipathos, banco de dados com diversas imagens de folhas saudáveis e doentes;
- OPENCV uma biblioteca para o desenvolvimento de Softwares na área de Visão Computacional.

#### 1.5 ESTRUTURA DO TRABALHO

O presente trabalho está estruturado em cinco capítulos. O primeiro capítulo serve como introdução, fornecendo uma visão geral do tema abordado e

apresentando os objetivos da pesquisa, ou seja, os resultados finais esperados desta investigação. Além disso, é fornecida a justificativa para a escolha do tema, demonstrando a relevância e a necessidade do estudo. Também são descritos os materiais e métodos utilizados ao longo do desenvolvimento do trabalho.

O segundo capítulo apresenta a fundamentação teórica adquirida por meio da revisão bibliográfica, abordando conceitos relacionados aos sistemas de detecção de doenças em folhas. Neste capítulo, são explorados os principais conhecimentos teóricos e técnicas utilizadas nessa área, fornecendo uma base sólida para o desenvolvimento do trabalho.

O terceiro capítulo aborda o desenvolvimento e a implementação do trabalho. Aqui, são detalhadas as etapas e os processos envolvidos na criação do sistema de detecção de doenças em folhas, desde a coleta e pré-processamento dos dados até a implementação das técnicas e algoritmos escolhidos.

No quarto capítulo, são apresentados e discutidos os resultados obtidos ao longo da pesquisa. São analisados os dados coletados, os resultados dos testes e experimentos realizados, bem como as observações e conclusões relevantes para o tema em questão.

Por fim, o quinto capítulo encerra o trabalho com a conclusão final, destacando as principais descobertas e insights obtidos por meio da investigação realizada. Também são apresentadas propostas de estudos para trabalhos futuros, visando a expansão e o aprimoramento das abordagens utilizadas no contexto de detecção de doenças em folhas.

A estrutura organizada desses capítulos permite a abordagem e sistemática do tema, garantindo uma compreensão do assunto e contribuindo para o avanço do conhecimento nessa área específica.

## 2 FUNDAMENTAÇÃO TEÓRICA E REVISÃO DE LITERATURA

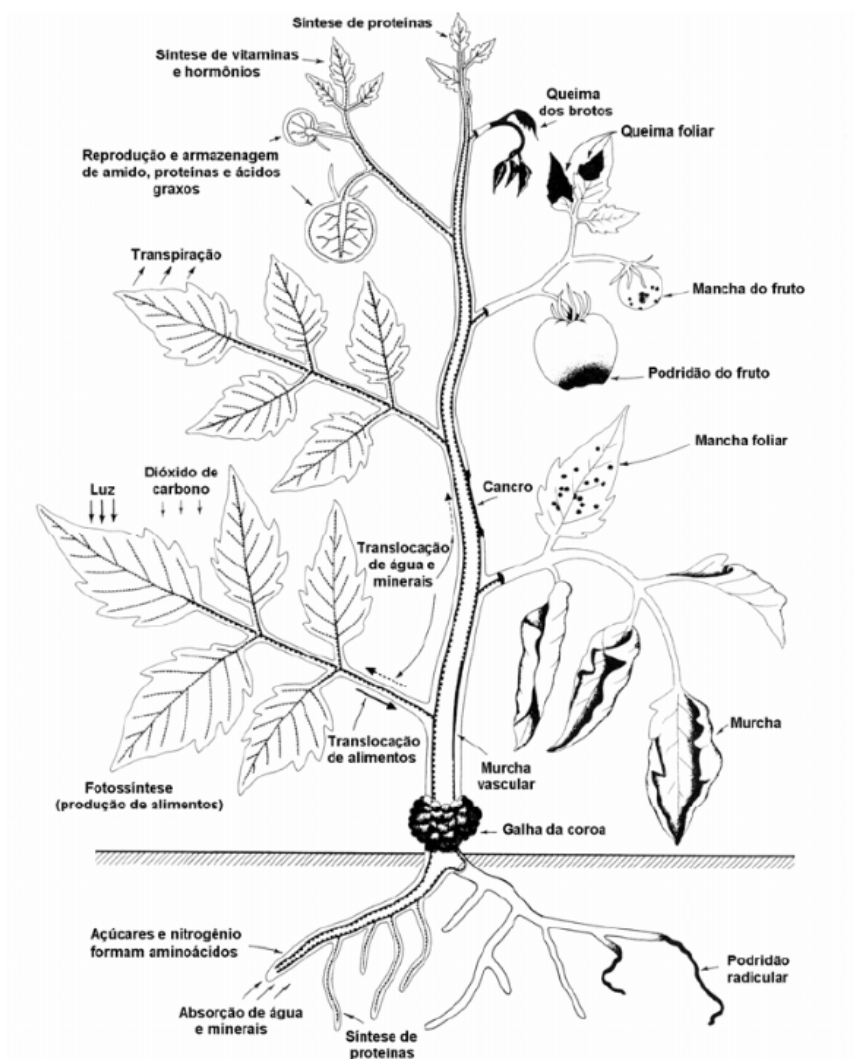
### 2.1 DOENÇA DE PLANTAS

As doenças em plantas são alterações no seu desenvolvimento fisiológico ou da estrutura das mesmas, causadas por fatores bióticos ou abióticos, que afetam negativamente a sua produtividade e qualidade (RAVAL e SARTORATO, 1994). Os fatores bióticos são organismos vivos que se alimentam de tecidos vivos (parasitas) ou mortos (saprófitas), como fungos, bactérias, vírus, nematóides, fitoplasmas, protozoários e plantas parasitas. Os fatores abióticos são condições ambientais desfavoráveis, como poluentes atmosféricos, desequilíbrios de nutrientes e estresses hídricos, térmicos ou luminosos (MICHEREFF, 2001).

Para ilustrar esses conceitos, na Figura 1 a seguir observa-se as diferentes partes da planta, como raízes, caule, folhas e flores, que desempenham funções essenciais, como absorção de água e nutrientes, transporte de seiva, fotossíntese e reprodução. Além disso, são mostrados sintomas característicos de doenças, como manchas nas folhas, murcha, deformação, necrose e presença de estruturas fúngicas.



Figura 1 – Representação esquemática das funções básicas da planta e sintomas causados por alguns tipos de doenças



Fonte: MICHEREFF(2001)

As doenças em plantas podem ser classificadas de acordo com a sua distribuição geográfica (endêmicas ou epidêmicas), a sua duração (perenes ou anuais), o seu agente causal (bióticas ou abióticas) e o seu modo de transmissão (contagiosas ou não contagiosas) (RAVAL e SARTORATO, 1994). A etiologia e a classificação dos agentes causais dependem das características morfológicas, fisiológicas e genéticas dos mesmos, bem como das suas relações com os hospedeiros e o ambiente.

A sintomatologia é o estudo dos sintomas das doenças em plantas, que são

as manifestações externas ou internas da alteração do seu desenvolvimento normal. Os sintomas podem ser causados diretamente pelo agente causal ou indiretamente pela reação da planta ao mesmo. Alguns exemplos de sintomas são: clorose, necrose, murcha, podridão, mancha foliar, ferrugem, oídio, cancro, galha e mosaico (MICHEREFF, 2001).

A epidemiologia é o estudo do desenvolvimento das doenças em plantas em uma população, em função do tempo e das condições ambientais. A interação entre o hospedeiro, o patógeno e o ambiente é denominada triângulo da doença. A epidemiologia visa entender os fatores que influenciam na incidência, na severidade e na dispersão das doenças em plantas, bem como prever e quantificar as suas perdas (MICHEREFF, 2001).

O controle das doenças em plantas é o conjunto de medidas que visam reduzir a sua incidência e severidade a níveis economicamente aceitáveis. O controle pode ser baseado em métodos genéticos (uso de cultivares resistentes), culturais (rotação de culturas, adubação equilibrada, irrigação adequada etc.), biológicos (uso de antagonistas naturais ou introduzidos), físicos (uso de calor, radiação ou eletricidade) ou químicos (uso de fungicidas, bactericidas ou nematocidas) (MICHEREFF, 2001).

## 2.2 DETECÇÃO DE DOENÇAS EM FOLHAS

A detecção de doenças em folhas de plantas é uma atividade importante para o manejo fitossanitário e a produtividade agrícola. As doenças em folhas podem causar redução na fotossíntese, na qualidade e na quantidade dos produtos agrícolas, além de facilitar a entrada de outros patógenos e pragas (MICHEREFF, 2001). As doenças em folhas são caracterizadas por sintomas visuais, como manchas, cloroses, necroses, verrugoses, ferrugens, oídios, cancos, galhas e mosaicos (EMBRAPA ALGODÃO, 2003).

A detecção automática de doenças em folhas pode proporcionar melhor monitoramento de grandes áreas agrícolas, reduzindo os custos e o tempo do exame visual tradicional realizado por especialistas. A detecção automática consiste em identificar e classificar as regiões de doenças em imagens digitais de folhas de plantas, utilizando técnicas de processamento digital de imagens e aprendizado de máquina (ACADEMIA EDU, 2016).

Uma das técnicas mais utilizadas para a detecção automática de doenças em folhas é o uso de redes neurais convolucionais (CNNs), que são modelos computacionais inspirados no funcionamento do córtex visual dos animais. As CNNs são capazes de extrair características relevantes das imagens e aprender a classificar as regiões de doenças em diferentes categorias, com alta precisão e robustez (USP, 2022).

Uma das abordagens mais recentes para a detecção automática de doenças em folhas é o uso de dois estágios independentes: um para realizar a detecção das regiões de doenças nas imagens e outro para realizar a classificação das mesmas. O primeiro estágio utiliza uma CNN para gerar candidatos a regiões de doenças, baseado em mapas de ativação da rede. O segundo estágio utiliza outra CNN para classificar os candidatos gerados pelo primeiro estágio, utilizando uma função de perda ponderada que leva em conta a confiança da detecção e a acurácia da classificação (USP, 2022).

Essa abordagem tem como vantagens a redução do número de falsos positivos na detecção, a melhoria da acurácia na classificação e a flexibilidade para lidar com diferentes tipos e graus de doenças em folhas. Além disso, essa abordagem pode ser otimizada por meio de algoritmos genéticos, que são técnicas inspiradas na evolução natural dos seres vivos. Os algoritmos genéticos podem ser utilizados para ajustar os parâmetros das redes neurais convolucionais, como o número e o tamanho dos filtros, as taxas de aprendizado e os pesos das funções de perda (ACADEMIA EDU, 2016).

A detecção automática de doenças em folhas é uma área promissora para o desenvolvimento de sistemas inteligentes que auxiliem os produtores rurais na tomada de decisão sobre o manejo fitossanitário das culturas. Esses sistemas podem contribuir para a redução das perdas causadas pelas doenças em folhas e para o aumento da sustentabilidade da produção agrícola.

### 2.3 REDES NEURAIIS ARTIFICIAIS

As redes neurais artificiais (RNAs) são modelos matemáticos que se inspiram nas estruturas neurais biológicas e adquirem capacidade computacional por meio do aprendizado e generalização (REZENDE, 2003). Compostas por neurônios artificiais interconectados por pesos sinápticos, as RNAs transmitem sinais e processam

informações. Cada neurônio artificial recebe um conjunto de entradas, pondera-as pelos pesos sinápticos, soma-as e aplica uma função de ativação para gerar uma saída (CAPES, 2019).

As RNAs podem ser organizadas em diferentes arquiteturas, dependendo da forma como os neurônios se conectam e se comunicam. As arquiteturas mais comuns são: redes feedforward, redes recorrentes e redes auto-organizáveis. Redes feedforward propagam os sinais em um único sentido, da camada de entrada à camada de saída, com camadas intermediárias ou ocultas. Redes recorrentes permitem a propagação dos sinais em ambos os sentidos, com conexões de realimentação. Redes auto-organizáveis ajustam os neurônios de acordo com as características dos dados de entrada, formando grupos ou mapas (IBM, 2021).

Na Figura 2 se observar o modelo matemático de um neurônio, as principais partes são:

- Entradas (X):

As entradas representam os valores numéricos fornecidos ao neurônio. Elas podem representar características, atributos ou informações relevantes para a tarefa em questão. Cada entrada é multiplicada pelo peso sináptico correspondente antes de ser somada.

- Pesos sinápticos (W):

Os pesos sinápticos são parâmetros ajustáveis que controlam a importância relativa de cada entrada na saída do neurônio. Cada entrada tem um peso associado que determina a contribuição daquela entrada específica na saída do neurônio. Os pesos sinápticos são ajustados durante o treinamento da rede neural para otimizar o desempenho do modelo.

- Soma ponderada ( $W \cdot X$ ):

A soma ponderada é o resultado da multiplicação de cada entrada pelo seu respectivo peso sináptico. Ela representa a combinação linear das entradas ponderadas pelos pesos sinápticos. A soma ponderada é um passo importante para determinar a influência de cada entrada na saída do neurônio.

- Viés (b):

O viés (bias) é um parâmetro adicional no modelo do neurônio artificial. Ele é um valor constante que é adicionado à soma ponderada das entradas e pesos

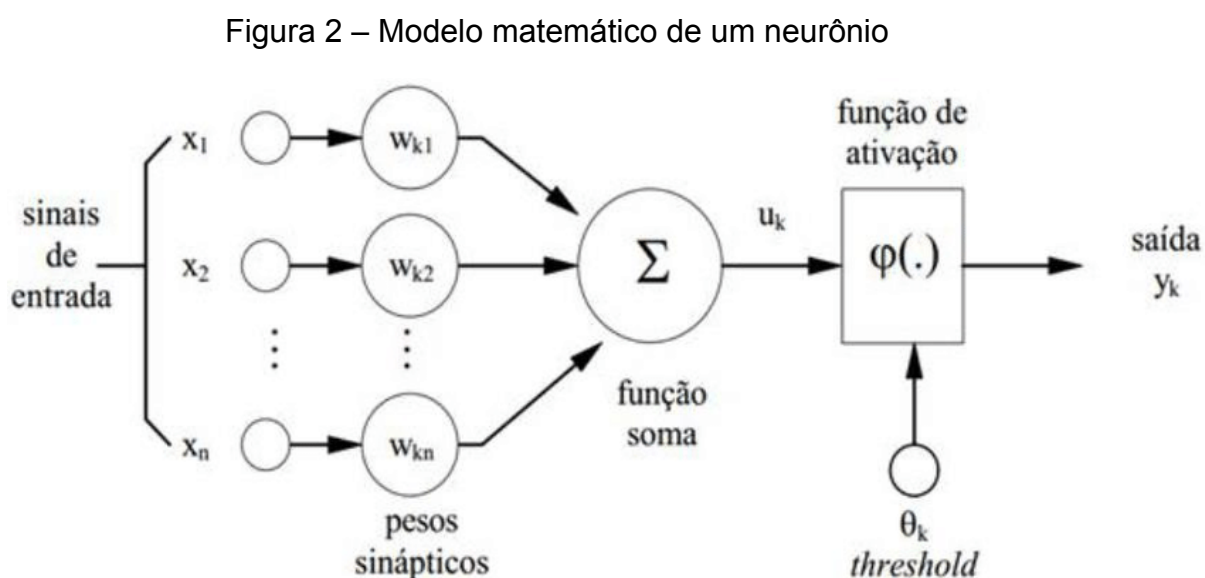
sinápticos. O viés permite ajustar o ponto de partida ou o limiar de ativação do neurônio, afetando a sensibilidade do neurônio a diferentes intervalos de entrada.

- Função de ativação (f):

A função de ativação é aplicada ao resultado da soma ponderada mais o viés. Ela introduz não-linearidades no modelo, permitindo que o neurônio aprenda a representar relações complexas entre as entradas e a saída. Existem várias funções de ativação comumente utilizadas, como a função sigmoide, função degrau, função ReLU (Rectified Linear Unit) e função tangente hiperbólica. Cada função de ativação tem características diferentes e é escolhida com base no tipo de problema e nas propriedades desejadas da rede neural.

- Saída (y):

A saída do neurônio é o resultado final após a aplicação da função de ativação. Ela representa o valor calculado ou a resposta do neurônio para uma determinada entrada. A saída do neurônio pode ser usada como entrada para outros neurônios em uma rede neural, formando uma estrutura de processamento em camadas.



Fonte: Adaptado de Haykin (2001)

Para treinar as RNAs, são utilizados diferentes algoritmos de aprendizagem, que ajustam os pesos sinápticos para minimizar o erro entre a saída desejada e a

saída gerada pela rede. Os algoritmos de aprendizado podem ser supervisionados, não supervisionados ou por reforço. Os supervisionados usam dados rotulados, com saídas desejadas conhecidas. Os não supervisionados usam dados não rotulados, sem saídas desejadas. Os algoritmos por reforço usam sinais de recompensa ou punição baseados no desempenho da rede (MEDIUM, 2018).

As RNAs têm diversas aplicações em reconhecimento de padrões, classificação, regressão, previsão, otimização, controle e inteligência artificial. Algumas vantagens das RNAs incluem a capacidade de lidar com dados ruidosos ou incompletos, adaptação a mudanças nos dados ou ambiente, generalização para novos casos e solução de problemas complexos e não lineares.

## 2.4 VISÃO COMPUTACIONAL

A visão computacional é uma área da inteligência artificial que tem como objetivo replicar as funções da visão humana em máquinas. Isso permite que elas sejam capazes de identificar e classificar objetos, pessoas, cenas e atividades em imagens e vídeos, e reagir de acordo com o que "veem" (SAS, 2021). A visão computacional busca compreender o conteúdo de vídeos e imagens estáticas, extrair informações úteis e aplicá-las para solucionar problemas em diversos domínios (NVIDIA, 2021).

Essa área apresenta desafios, pois envolve o processamento de grandes volumes de dados visuais, a representação de características relevantes das imagens, a modelagem das relações entre os elementos visuais e o contexto, a inferência de conhecimento a partir das imagens e a integração com outras modalidades sensoriais (PIXFORCE, 2020).

Para realizar suas tarefas, a visão computacional utiliza métodos de processamento digital de imagens e aprendizado de máquina. Isso inclui técnicas como segmentação, classificação, detecção, reconhecimento, restauração e geração de imagens. Um dos métodos mais utilizados atualmente são as redes neurais convolucionais (CNNs), que são modelos computacionais inspirados no funcionamento do córtex visual dos animais. As CNNs são capazes de extrair características hierárquicas das imagens e aprender a classificar as regiões de interesse em diferentes categorias com alta precisão e robustez (SAS, 2021).

A visão computacional possui inúmeras aplicações em diversos setores,

incluindo esportes, automóveis, agricultura, varejo, bancos, construção, seguradoras e saúde. Alguns exemplos de aplicação são: reconhecimento facial, detecção de placas de veículos, análise de tráfego, monitoramento de segurança, diagnóstico médico por imagem, detecção de doenças em plantas, realidade aumentada e virtual, entretenimento e arte (CRM PIPE RUN, 2021).

A visão computacional é uma área promissora para o desenvolvimento de sistemas inteligentes que auxiliem os seres humanos na tomada de decisões sobre problemas do mundo real. Esses sistemas podem contribuir para a melhoria da qualidade de vida, produtividade e sustentabilidade da sociedade como um todo.

## 2.5 IMAGEM DIGITAL

A imagem digital, uma forma inovadora de expressar objetos, cenas e fenômenos, é materializada por meio de dados numéricos. Esses dados podem ser armazenados, transmitidos, impressos ou processados por dispositivos eletrônicos. A gênese de uma imagem digital é variada, podendo advir de câmeras fotográficas, scanners, softwares de desenho, simulação, entre outros. Com aplicações abrangentes, a imagem digital é empregada em campos tão diversos quanto arte, comunicação, educação, entretenimento, engenharia e medicina (GONZALEZ, 2010).

Um retrato digital é composto por uma finita coleção de pontos denominados pixels (*picture elements*), sendo eles os menores componentes visíveis da imagem. Cada pixel ostenta uma localização e uma cor codificada, traduzida em números binários (0 ou 1). A congregação de vários bits viabiliza a representação de diferentes graus de intensidade ou matizes de cor para cada pixel. Por exemplo, caso um pixel seja representado por 8 bits, terá 256 opções distintas, equivalentes a 256 tons de cinza. Caso esse valor se estenda para 24 bits, o pixel comportará mais de 16 milhões de possibilidades, abarcando uma ampla gama de cores (GONZALEZ, 2010).

A qualidade de uma imagem digital repousa sobre dois fatores cruciais: a resolução e as dimensões. A resolução, medida em pixels por polegada (PPI) ou pontos por polegada (DPI), quantifica a quantidade de pixels por unidade de área da imagem. Uma resolução mais alta promove mais detalhes e nitidez na imagem. Já as dimensões expressam a quantidade total de pixels na imagem, podendo ser

aferidas em termos de largura, altura (em pixels) ou área (em megapixels). Dimensões maiores incorporam maior quantidade de informação.

No âmbito das imagens digitais, são identificados dois tipos fundamentais: a imagem raster e a imagem vetorial. A imagem raster retrata cada pixel de forma individual, constituindo uma matriz ou grelha de pontos. Já a imagem vetorial expressa objetos por meio de equações matemáticas, traduzindo-se em curvas e formas geométricas. A imagem raster se sobressai em imagens fotográficas ou realistas, apta a reproduzir sutilezas e sombras com exatidão. Por sua vez, a imagem vetorial se revela mais indicada para imagens gráficas ou ilustrativas, possibilitando a modificação e redimensionamento dos objetos sem perda de qualidade.

A imagem digital floresce como campo de estudo e pesquisa, abarcando uma profusão de disciplinas, como matemática, computação, física, biologia e arte. Os avanços nessa esfera impulsionam o desenvolvimento de técnicas e tecnologias que elevam a capacidade humana de capturar, transmitir, manipular e interpretar imagens digitais. Enquanto ocorre uma evolução constante, surge uma compreensão cada vez mais profunda da intersecção entre a representação visual e o mundo dos dados numéricos.

### **2.5.1 Sistemas de Cores**

Um sistema de cores é um conjunto de diretrizes e convenções que possibilitam a representação e manipulação de cores em diversos meios, tais como luz, tinta, papel, monitores, entre outros. A variedade de sistemas de cores disponíveis oferece características singulares, acompanhadas por vantagens e desvantagens distintas. Neste estudo, será focado em duas abordagens amplamente empregadas na computação gráfica: o sistema RGB e o sistema HSV.

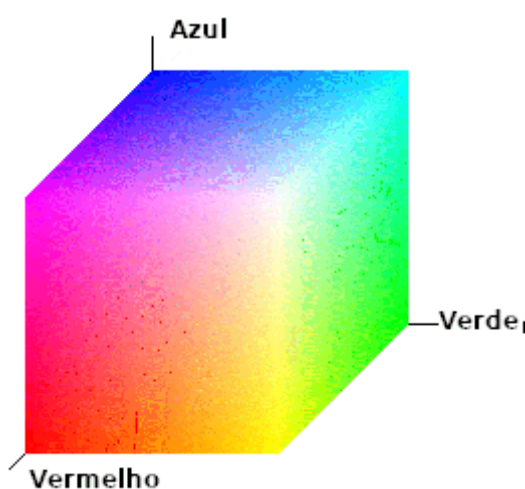
#### **2.5.1.1 Sistema de Cores RGB**

O sistema RGB, um modelo aditivo de cores, é construído pela combinação das luzes vermelha, verde e azul (*Red*, *Green* e *Blue* em inglês). Cada uma dessas cores primárias pode ter uma intensidade variável, geralmente expressa em valores entre 0 e 255. Desta forma, uma cor no sistema RGB é definida por um conjunto de



três valores, exemplificado por (255, 0, 0) para o vermelho, (0, 255, 0) para o verde e (0, 0, 255) para o azul. O resultado da combinação dessas cores primárias em proporções diferentes resulta em uma variedade de outras cores. Por exemplo, o amarelo é expresso como (255, 255, 0), o magenta como (255, 0, 255) e o ciano como (0, 255, 255). A mescla equitativa das três cores primárias gera o branco (255, 255, 255), enquanto a ausência total delas resulta no preto (0, 0, 0). As cores e suas combinações são mostradas na Figura 3.

Figura 3 – Sistema de Cores RGB



Fonte: Adaptado de MICROSOFT

Amplamente empregado em dispositivos que emitem luz, como monitores, televisores, projetores e câmeras digitais, o sistema RGB é capaz de reproduzir uma vasta gama de cores, cerca de 16 milhões. No entanto, não engloba todas as nuances perceptíveis ao olho humano. Além disso, a representação RGB pode variar conforme as especificações do dispositivo e as condições do ambiente. Dessa forma, a calibração adequada dos dispositivos é essencial para preservar a fidelidade das cores.

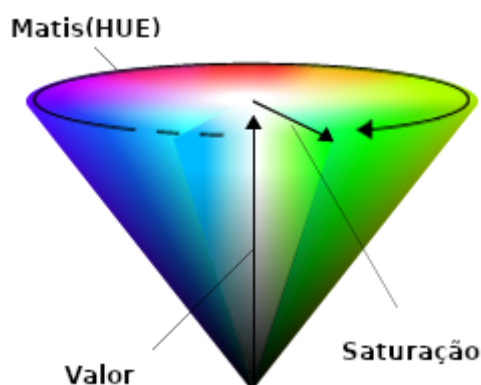
A simplicidade e a facilidade de implementação do sistema RGB em dispositivos eletrônicos são vantagens notáveis. No entanto, o sistema apresenta uma desconexão com a forma como os seres humanos percebem as cores. Determinar a luminosidade relativa de uma cor ou ajustar sua tonalidade e saturação exclusivamente com base nos valores RGB pode ser desafiador.

Conforme Ferreira e Spricigo (2017), o sistema RGB se baseia na combinação de intensidades variáveis das cores vermelha, verde e azul. Eles destacam sua ampla adoção na computação gráfica devido à compatibilidade com dispositivos emissores de luz. No entanto, ressaltam as limitações do sistema para representar todas as cores discerníveis pelo olho.

#### 2.5.1.2 Sistema de Cores HSV

O sistema HSV, em contraposição ao RGB, almeja uma representação das cores mais alinhada à percepção humana. Ele compreende as componentes *hue* (matiz), *saturation* (saturação) e *value* (valor). A matiz define a cor pura do espectro visível, variando de 0 a 360 graus, correspondendo às cores do arco-íris: vermelho (0°), laranja (30°), amarelo (60°), verde (120°), ciano (180°), azul (240°), magenta (300°) e vermelho novamente (360°). A saturação determina a pureza ou intensidade da cor, variando entre 0 e 100%, onde 0% indica um tom acinzentado e 100% denota saturação total. O valor define o brilho ou luminosidade da cor, também entre 0 e 100%, com 0% representando preto e 100% representando branco. Se observa na Figura 4 o cone que representa o sistema.

Figura 4 – Sistema de Cores HSV



Fonte: Adaptado de MICROSOFT

Amplamente empregado em edição de imagens e design gráfico, o sistema HSV permite uma manipulação mais intuitiva e criativa das cores. Alterar a tonalidade ou saturação sem afetar o brilho é uma das possibilidades oferecidas.

Além disso, a ordenação das cores por matiz ou valor possibilita a criação de paletas harmônicas ou contrastantes.

Uma das principais vantagens do sistema HSV reside em sua maior congruência com a percepção humana das cores. Porém, sua complexidade e a dificuldade de implementação em dispositivos eletrônicos são desafios notáveis. Por essa razão, é comum converter as cores do sistema HSV para o sistema RGB antes de exibi-las ou armazená-las.

### 2.5.2 HOG

O Histograma de Gradientes Orientados (HOG), também conhecido como *Histogram of Oriented Gradients*, é um descritor de características amplamente utilizado na extração de informações de imagens, especialmente em tarefas de visão computacional voltadas para a detecção de objetos.

Uma das características distintivas do HOG é seu foco na estrutura interna do objeto em vez de sua aparência visual. Ele se baseia na análise das magnitudes e orientações dos gradientes presentes em cada célula da imagem. Esse enfoque permite uma compreensão mais profunda da composição estrutural dos objetos e das formas de contorno que eles apresentam.

O processo do HOG começa com a conversão da imagem para uma janela de detecção de tamanho fixo, normalmente 64x128 pixels. A partir dessa forma, a imagem é subdividida em células menores, geralmente de dimensões 8x16 pixels. Em seguida, são calculados os gradientes e suas orientações para cada célula, constituindo os componentes fundamentais das características extraídas.

Outro aspecto essencial do HOG é a divisão da imagem em blocos que englobam um número definido de células. Esses blocos podem apresentar uma sobreposição de 50%. Cada bloco contém um conjunto de células, e para cada bloco é criado um histograma com um número específico de "*bins*". Esses histogramas refletem a distribuição dos gradientes orientados nas células do bloco. O resultado desse processo é um vetor de características consolidado, composto pelos histogramas normalizados de todos os blocos na imagem.

O HOG oferece várias vantagens distintas em relação a outros descritores de características. Ele demonstra robustez em face de variações de iluminação, escala

e rotação, graças à normalização dos gradientes e histogramas. Além disso, é capaz de capturar informações valiosas sobre as formas e contornos dos objetos, uma vez que se baseia nas orientações dos gradientes. Sua implementação é relativamente simples, utilizando operações matemáticas básicas, como o cálculo de gradientes e a criação de histogramas (Singh, 2023).

No entanto, o HOG também apresenta algumas limitações. Ele pode ser sensível a deformações não-rígidas em objetos, que são alterações na forma que não mantêm a relação espacial entre os pontos, uma vez que utiliza uma janela de detecção fixa. Além disso, o processo de extração de características do HOG pode gerar vetores de dimensões consideráveis, devido à inclusão de múltiplos blocos e células na imagem. Pode-se, também, ter dificuldade em distinguir objetos com formas semelhantes ou texturas complexas, pois sua abordagem se baseia predominantemente em gradientes locais (Singh, 2023).

Em resumo, o HOG é um descritor de características poderoso e versátil, com ampla aplicação na detecção de objetos em imagens. Sua capacidade de focar a estrutura e os contornos dos objetos, juntamente com sua robustez a variações de iluminação e escala, o tornam uma ferramenta valiosa em tarefas de visão computacional. No entanto, é importante considerar suas limitações e adequar sua utilização às necessidades específicas de cada aplicação.

### **2.5.3 Segmentação de Imagens**

A segmentação de imagem é um processo fundamental na análise de dados visuais, visando a divisão da imagem em regiões ou objetos baseada em critérios de similaridade ou homogeneidade. Seu objetivo reside em simplificar ou remodelar a representação da imagem, tornando-a mais acessível para análise ou extração de informações relevantes. Essa técnica desempenha um papel central em diversas aplicações de visão computacional, abrangendo desde o reconhecimento de objetos e detecção de faces até análises médicas e processamento de imagens de satélite (GONZALEZ, 2010).

No cenário da segmentação de imagem, uma série de métodos e algoritmos estão disponíveis, sendo classificados em distintas categorias conforme o tipo de informação empregada, o grau de interação do usuário, o domínio da imagem e o resultado desejado. Dentre esses métodos, destacam-se alguns exemplos:

1. Baseados em Regiões: Os métodos baseados em regiões agrupam pixels com características semelhantes, tais como cor, intensidade, textura ou forma. Isso se desdobra em dois tipos principais: crescimento de regiões (*region growing*) e divisão e fusão (*split and merge*). O primeiro parte de sementes iniciais (*seeds*), que podem ser definidas pelo usuário ou determinadas automaticamente, expandindo regiões vizinhas que atendem a um critério de similaridade. No segundo, a imagem é segmentada em sub-regiões menores e homogêneas, posteriormente combinadas caso sejam similares.
2. Baseados em Bordas: Essa abordagem visa identificar descontinuidades ou transições abruptas nos valores dos pixels, que indicam as fronteiras entre objetos ou regiões. O emprego de operadores diferenciais, como gradientes ou laplacianos, ou filtros específicos, como Sobel ou Canny, realça as bordas. Posteriormente, técnicas de limiarização, afinamento ou ligação são aplicadas para obter os contornos dos objetos.
3. Baseados em Histograma: A análise da distribuição dos valores dos pixels, conhecida como histograma, é o cerne desse método. Picos ou modos no histograma podem indicar diferentes classes ou objetos. Técnicas estatísticas, como análise de agrupamentos ou mistura de gaussianas, são usadas para identificar os modos e atribuir pixels às classes correspondentes. Dentro do método baseado em histograma, uma das abordagens mais simples e amplamente utilizadas é o método de Limiarização (Thresholding). Nesse método, um valor de limiar é escolhido, e todos os pixels com valores abaixo desse limiar são atribuídos a uma classe, enquanto os pixels com valores acima do limiar são atribuídos a outra classe. Embora seja uma estratégia ágil e simples, a escolha do limiar correto é crucial, sendo suscetível a ruídos e variações na iluminação.
4. Baseados em Grafos: Aqui, a imagem é representada como um grafo, no qual os pixels ou regiões são vértices e as relações entre eles, como distância ou similaridade, são arestas. Algoritmos de

particionamento de grafos, como corte mínimo ou corte normalizado, são aplicados para dividir o grafo em subgrafos correspondentes a classes ou objetos.

A seleção do método de segmentação mais adequado depende da natureza da imagem, dos objetivos da segmentação e dos recursos disponíveis. Avaliar a qualidade e eficiência da segmentação obtida é igualmente importante, com métricas quantitativas e qualitativas servindo como ferramentas de avaliação. Métricas quantitativas podem comparar a segmentação com um padrão-ouro ou considerar propriedades intrínsecas da segmentação. Métricas qualitativas, por sua vez, ponderam a percepção humana e fatores subjetivos (GONZALEZ, 2010).

#### **2.5.4 Processamento de Imagens**

A área do processamento de imagem, inserida no âmbito da ciência da computação, dedica-se de maneira minuciosa à manipulação e análise de dados visuais, os quais englobam fotografias, vídeos e imagens digitais. Seu propósito reside em extrair informações de relevância, ampliar a qualidade, estabelecer novas representações ou realizar operações específicas sobre as imagens em questão. Neste escrito, é empreendida uma revisão bibliográfica pormenorizada acerca do processamento de imagem, explorando-se seus fundamentos, metodologias, aplicações e complexidades subjacentes (GONZALEZ, 2010).

O processo de processamento de imagem pode ser destrinchado em três etapas primordiais: aquisição, processamento e análise. A fase de aquisição engloba a captura das imagens por intermédio de sensores ópticos, a exemplo de câmeras, scanners e satélites. O processo de processamento propriamente dito envolve a aplicação de técnicas matemáticas e algorítmicas visando a modificações nas imagens, como filtragem, segmentação, realce, compressão ou transformação. A análise das imagens compreende a tarefa de interpretação, abrangendo desde o reconhecimento de padrões até a detecção de arestas, extração de características e classificação (GONZALEZ, 2010).

O campo do processamento de imagem desdobra-se em uma diversidade de aplicações que abraçam distintos domínios do conhecimento, tais como medicina, engenharia, astronomia, biologia, geografia, segurança, arte e entretenimento. Algumas das aplicações de destaque englobam o diagnóstico médico por meio de

imagens, a restauração de registros visuais históricos, a detecção de objetos em imagens aéreas ou espaciais, além do reconhecimento facial e de impressões digitais, não obstante a criação de efeitos especiais e realidade aumentada.

No entanto, o processamento de imagem também se depara com uma série de desafios e limitações, intrinsecamente ligados à qualidade das imagens, ao escopo do problema e ao método adotado. Algumas das adversidades incluem a necessidade de lidar com ruídos, distorções ou oclusões presentes nas imagens, a seleção criteriosa de parâmetros específicos para cada técnica empregada, a avaliação rigorosa dos resultados obtidos, bem como a integração harmoniosa de diversas fontes de informação. Além disso, a consideração de questões éticas e de privacidade é fundamental no emprego responsável das imagens no campo do processamento.

## 2.6 TRABALHOS CORRELATOS

Nos últimos anos, tem havido uma crescente quantidade de abordagens em visão computacional e processamento de imagem que foram propostas para resolver desafios relacionados à agricultura (FERREIRA, 2017). Pesquisas recentes destacam a viabilidade da utilização de imagens para identificar e classificar fitopatógenos em várias culturas, com base em imagens de folhas.

Barbedo (2015) desenvolveu um algoritmo fundamentado em técnicas tradicionais de processamento digital de imagens, combinado com um sistema de votação em pares. Esse algoritmo produz uma lista de doenças, juntamente com suas respectivas probabilidades de ocorrência na folha analisada. A abordagem baseia-se principalmente nas informações de cor, onde o espaço de cores RGB original é transformado em espaços de cores HSV, LAB e CMYK. Posteriormente, histogramas de intensidade são extraídos das representações em escala de cinza de cada um dos dez canais resultantes.

Pires et al. (2016) propuseram uma abordagem que combinou o uso de descritores locais de imagem com a técnica "*Bag of Visual Words*" em conjunto com *Support Vector Machines* (SVM). As SVM são um tipo de algoritmo de aprendizado de máquina que é usado para tarefas de classificação e regressão. Elas funcionam encontrando um hiperplano de decisão que melhor separa as classes de dados no espaço multidimensional. Em outras palavras, o SVM é capaz de identificar

fronteiras de decisão ótimas para separar diferentes categorias de dados. Neste contexto, o SVM foi utilizado para classificar amostras de folhas de soja como saudáveis ou afetadas por doenças com base nos descritores locais extraídos.

A técnica "*Bag of Visual Words*" é uma abordagem comum em visão computacional e processamento de imagem que se assemelha ao conceito de um dicionário visual. Ela envolve a criação de um conjunto de palavras visuais (também chamadas de "*visual words*" ou "vocabulário") a partir dos descritores locais de imagens. Esses descritores são extraídos de regiões específicas da imagem e usados para criar um histograma que representa a frequência de ocorrência de cada palavra visual no conjunto de dados. Isso permite que as características distintivas das imagens sejam representadas e comparadas de forma eficaz. Neste estudo, a técnica "*Bag of Visual Words*" foi empregada para criar representações compactas e informativas das folhas de soja com base em seus descritores locais.

Ao testarem essa abordagem em um conjunto de dados composto por 1.200 folhas de soja digitalizadas, abrangendo amostras saudáveis e amostras com evidências de três doenças comuns em lavouras de soja (Míldio, Ferrugem Tan e Ferrugem RB), eles alcançaram uma precisão de aproximadamente 98% para as doenças analisadas.

Abordagens mais recentes, como a de Fuentes et al. (2017), introduzem o uso de aprendizado profundo para detecção de doenças e pragas em plantas de tomateiro. Essas abordagens utilizam meta-arquiteturas de Redes Neurais Convolucionais (CNNs), que desempenham um papel fundamental no processamento de dados visuais complexos. As meta-arquiteturas são estruturas de alto nível que fornecem um esqueleto ou um conjunto de diretrizes para o design das redes neurais. Entre as meta-arquiteturas utilizadas, destacam-se o *Faster R-CNN* (*Region-based Convolutional Neural Network*), R-FCN (*Region-based Fully Convolutional Network*) e SSD (*Single Shot MultiBox Detector*).

O *Faster R-CNN* é uma meta-arquitetura que aprimora a detecção de objetos em imagens por meio da introdução de uma rede neural convolucional que propõe regiões candidatas de interesse, as chamadas "regiões de interesse" (ROIs), e em seguida, classifica essas ROIs para determinar a presença ou ausência de objetos de interesse, como doenças ou pragas em plantas de tomateiro.

Já o R-FCN é uma meta-arquitetura que utiliza uma abordagem



completamente convolucional para a detecção de objetos. Ele opera em toda a imagem de forma eficiente, eliminando a necessidade de gerar múltiplas propostas de regiões. Isso torna o R-FCN uma escolha eficaz para tarefas de detecção em tempo real, como a detecção de doenças em plantas de tomateiro.

Outra meta-arquitetura empregada é o SSD, que se destaca por sua capacidade de detecção de objetos em múltiplas escalas e tamanhos simultaneamente, tornando-o eficaz para detecção de doenças e pragas em plantas de tomateiro em diversas condições.

Além disso, essas meta-arquiteturas são frequentemente aprimoradas com o uso de "extratores de recursos profundos", como a VGGnet e a Residual Network (ResNet). Estes extratores desempenham um papel crucial na detecção de doenças e pragas em plantas de tomateiro, pois são responsáveis por capturar informações visuais discriminativas das imagens.

A VGGnet é notável por sua profundidade, o que significa que consiste em muitas camadas convolucionais empilhadas umas sobre as outras. Essa profundidade permite que a VGGnet aprenda características visuais complexas em diferentes níveis de abstração, tornando-a eficaz na representação de padrões sutis nas imagens de plantas de tomateiro. Essa capacidade de aprendizado de recursos visuais é essencial para identificar doenças e pragas com alta precisão.

Por outro lado, a Residual Network, ou ResNet, ganhou destaque por sua capacidade de treinar redes ainda mais profundas. A ResNet introduz blocos residuais que permitem que informações relevantes fluam mais facilmente através das camadas da rede, o que facilita o treinamento de redes profundas e, por consequência, melhora a capacidade de reconhecimento de padrões complexos nas imagens das plantas de tomateiro.

Assim, ao combinar essas meta-arquiteturas de detecção, como Faster R-CNN, R-FCN e SSD, com os extratores de recursos profundos VGGnet e ResNet, as abordagens alcançam resultados promissores, com precisões acima de 90%, e conseguem realizar a detecção de doenças e pragas em tempo real nas plantas de tomateiro.

Hassan et al. (2021) exploraram uma abordagem mais abrangente, empregando um amplo conjunto de imagens abrangendo diversas culturas. Eles utilizaram arquiteturas de CNN (Redes Neurais Convolucionais) mais compactas,

incluindo InceptionV3, InceptionResNetV2, MobileNetV2 e EfficientNetB0. É importante compreender o que essas arquiteturas representam:

**InceptionV3:** InceptionV3 é uma arquitetura de rede neural convolucional que se destaca por seu design de módulos de convolução chamados "módulos Inception". Esses módulos têm diferentes tamanhos de filtros convolucionais que permitem à rede extrair características em várias escalas. InceptionV3 é conhecida por sua eficiência e precisão em tarefas de classificação de imagens.

**InceptionResNetV2:** InceptionResNetV2 é uma extensão da arquitetura Inception que incorpora blocos residuais, semelhantes aos encontrados na ResNet. Isso ajuda a aprofundar a rede e a melhorar o fluxo de informações, resultando em um modelo mais eficaz para tarefas de visão computacional.

**MobileNetV2:** MobileNetV2 é uma arquitetura de rede neural convolucional projetada para eficiência computacional e tamanho reduzido. É adequada para aplicações em dispositivos móveis e sistemas embarcados, mantendo um desempenho sólido na tarefa de classificação de imagens. A sua eficiência é alcançada através do uso de camadas convolucionais especiais chamadas de "blocos de convolução separável".

**EfficientNetB0:** EfficientNetB0 é parte de uma família de arquiteturas de redes neurais convolucionais chamadas "EfficientNets". Essas redes foram projetadas para serem altamente eficientes em termos de uso de recursos computacionais, mas ainda assim alcançar um desempenho de ponta em tarefas de visão computacional. O número "B0" indica o tamanho-base da rede, sendo B0 a versão mais leve e B7 a versão mais pesada.

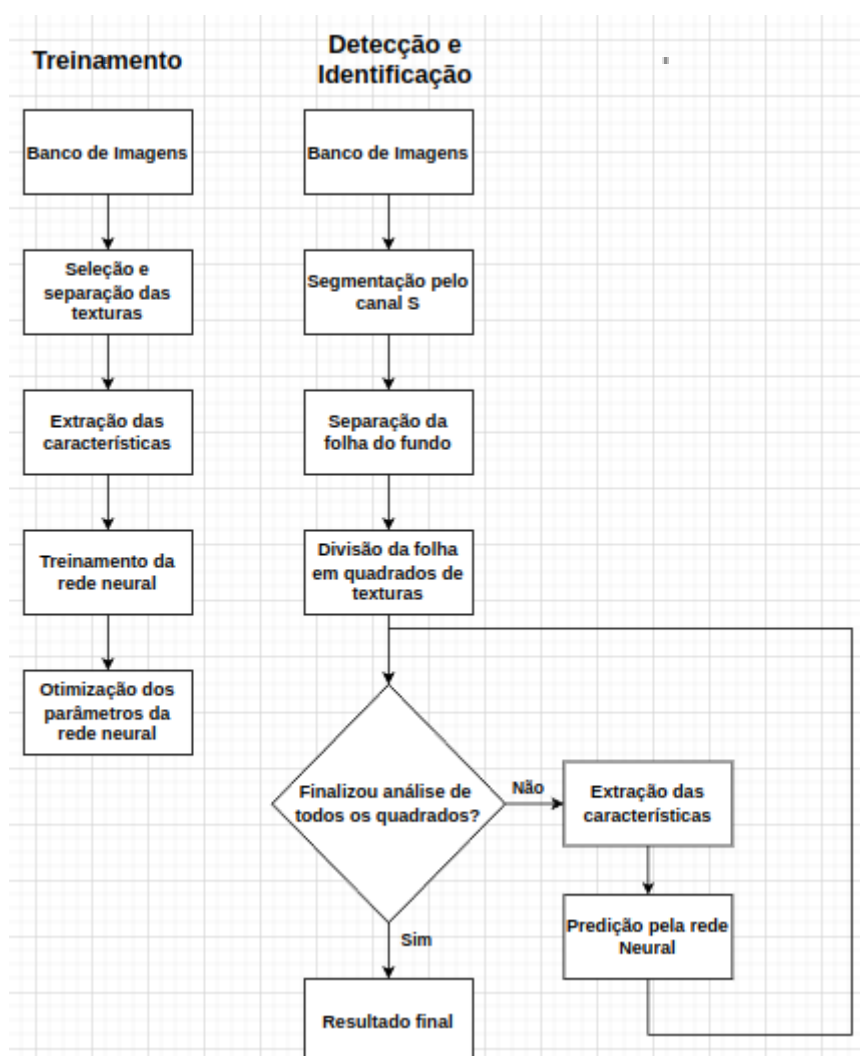
Os modelos treinados com essas arquiteturas apresentaram uma acurácia superior a 97%. É importante notar que esses resultados impressionantes foram obtidos com um conjunto de dados robusto, onde mesmo a classe com o menor tamanho continha 152 imagens, enquanto outras classes tinham mais de 5000 imagens. Isso demonstra a eficácia dessas arquiteturas em tarefas de classificação de imagens, especialmente quando se trata de um grande conjunto de dados e classes desequilibradas.

Essas abordagens ilustram a constante evolução das técnicas de visão computacional e processamento de imagem aplicadas à agricultura, oferecendo insights valiosos para a detecção e classificação de doenças em plantas.

### 3 DESENVOLVIMENTO DO TRABALHO

O diagrama de blocos apresentado na Figura 5 desempenha um papel fundamental ao ilustrar a arquitetura implementada. Este diagrama oferece uma visão abrangente das diferentes etapas e componentes do sistema, desde a aquisição das imagens das folhas até a fase de análise e diagnóstico. Cada bloco no diagrama representa um elemento-chave na cadeia de processamento, demonstrando a interconexão entre as partes e destacando a ordem lógica das operações realizadas.

Figura 5 – Diagrama de blocos do Sistema proposto



Fonte: o Autor

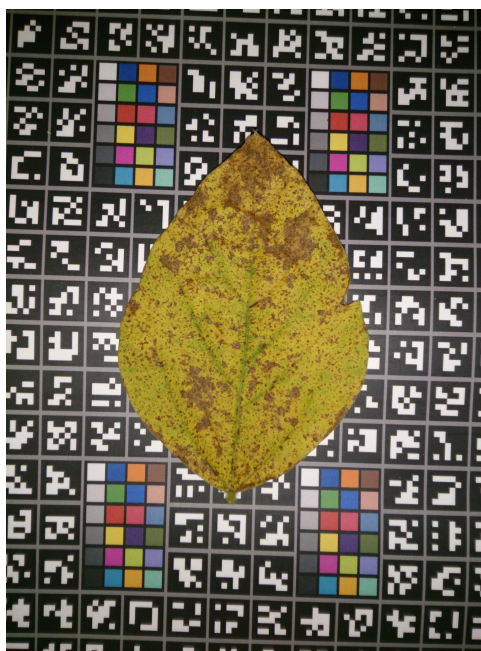
### 3.1 BANCO DE IMAGENS

Para o treinamento eficaz da rede neural proposta neste estudo, foi empregada uma base de imagens disponibilizada pela Embrapa chamada Digipathos. Essa base de dados consiste em 80 imagens que foram categorizadas em cinco classes distintas de doenças, além de uma categoria referente a folhas saudáveis.

É fundamental destacar a importância da escolha dessa base de dados no contexto deste estudo. As cinco classes de doenças incluídas são:

- Ferrugem: A ferrugem é uma doença fúngica que afeta as folhas das plantas, incluindo a soja. Ela se caracteriza por manchas de cor alaranjada a marrom nas folhas, como pode ser observado no exemplo presente na Figura 6.

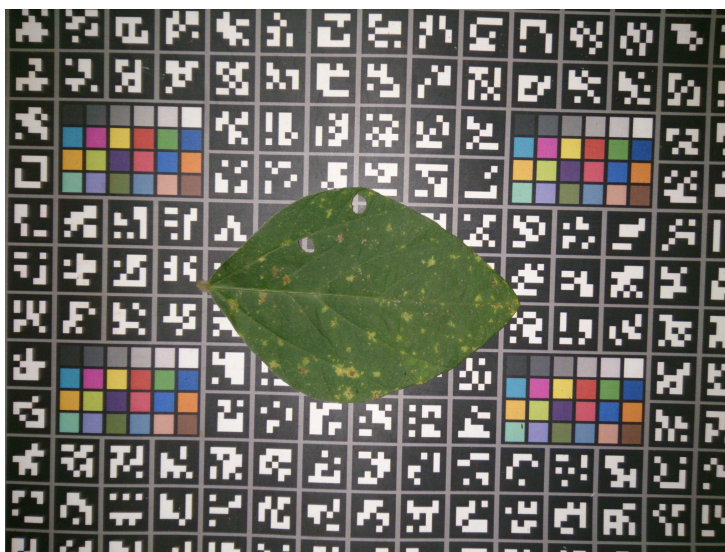
Figura 6 – Exemplar da doença Ferrugem



Fonte: EMBRAPA

- Míldio: O míldio é outra doença fúngica que pode afetar diversas culturas, incluindo a soja. Ele se manifesta como manchas escuras nas folhas, que podem se expandir rapidamente, como exemplificado na Figura 7.

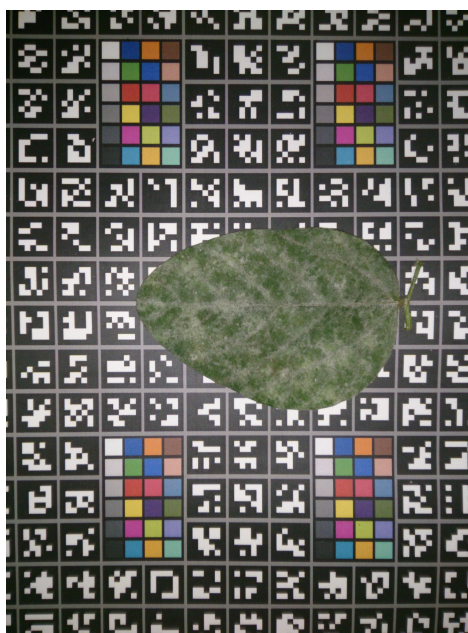
Figura 7 – Exemplar da doença Míldio



Fonte: EMBRAPA

- Oídio: O oídio é uma doença fúngica que se manifesta como um revestimento branco ou cinza nas folhas, assemelhando-se a uma camada de pó, como pode ser visto no exemplo na Figura 8.

Figura 8 – Exemplar da doença Oídio



Fonte: EMBRAPA

- Carijó: O carijo, também conhecido como mancha de carijó, é uma doença que se manifesta como pequenas manchas escuras ou marrons nas folhas, como ilustrado na Figura 9.

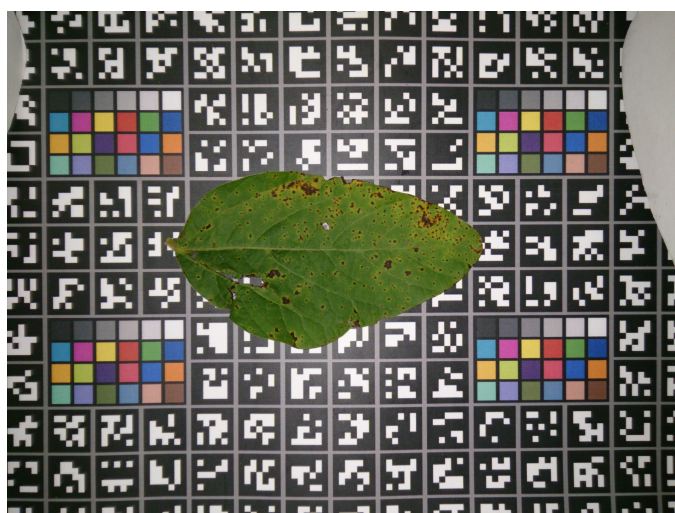
Figura 9 – Exemplar da doença Carijó



Fonte: EMBRAPA

- Crestamento: O crestamento, ou crestamento bacteriano, é uma doença bacteriana que causa manchas escuras, frequentemente com um centro necrótico, nas folhas, conforme exemplificado na Figura 10.

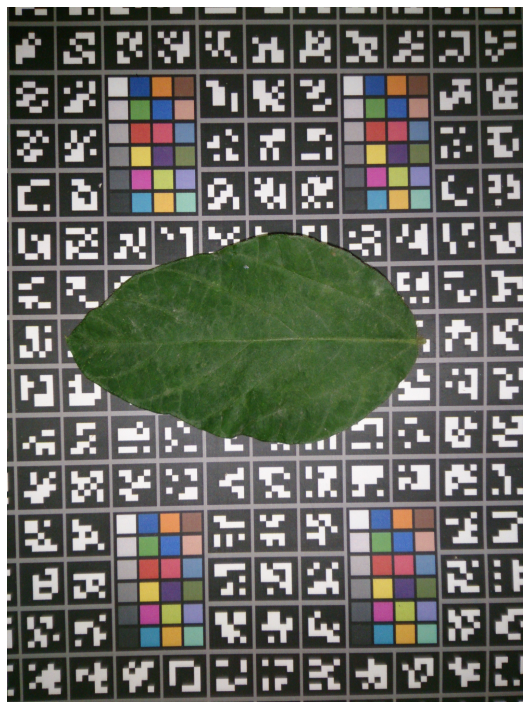
Figura 10 – Exemplar da doença Crestamento



Fonte: EMBRAPA

Além dessas cinco classes de doenças, uma categoria adicional foi reservada para representar folhas saudáveis, que não apresentam sintomas de nenhuma das doenças mencionadas, como ilustrado na Figura 11.

Figura 11 – Exemplar da folha saudável



Fonte: EMBRAPA

É importante ressaltar que a escolha desta base de dados é essencial para o desenvolvimento e avaliação do desempenho da rede neural no reconhecimento de doenças em folhas, proporcionando uma fundação sólida para a realização deste estudo.

### 3.2 SELEÇÃO DE IMAGENS

No processo de criação do banco de imagens, a extração de texturas das folhas desempenha um papel fundamental na caracterização das condições de saúde das plantas. Para esse propósito, texturas foram manualmente selecionadas a partir de áreas afetadas por doenças, correspondendo às folhas doentes, bem como texturas de áreas saudáveis, que representam as folhas em bom estado. Ao concentrar-se em texturas específicas de regiões doentes e saudáveis, o objetivo é

capturar as características distintas que podem ser exploradas pela rede neural para a identificação precisa de doenças.

Para enriquecer a base de imagens, foram extraídos diversos tipos de texturas, cujo número total pode ser consultado na Tabela 1. Essa abordagem visa equipar a rede neural com um conjunto diversificado e representativo de padrões texturais. Isso contribui significativamente para a capacidade de discriminação eficaz entre folhas doentes e saudáveis durante o processo de diagnóstico.

Tabela 1 – Número de texturas utilizadas por classe para treino

Classe da Folha	Número de <i>Frames</i>
Carijó	71
Ferrugem	93
Oídio	206
Crestamento	83
Míldio	57
Saudável	141

Fonte: O Autor

Além disso, como parte da padronização do conjunto de dados, todas as imagens foram redimensionadas para o tamanho uniforme de 64x64 pixels. Essa etapa garante que todas as imagens tenham a mesma dimensão, o que é crucial para que a rede neural processe as informações de maneira consistente e eficiente.

### 3.3 EXTRAÇÃO DAS CARACTERÍSTICAS

No processo de extração de características das texturas das folhas, foram empregadas abordagens fundamentais que desempenham um papel crucial na distinção entre folhas saudáveis e doentes. Uma dessas abordagens é a técnica HOG, que desempenha um papel crucial na captura de informações relacionadas aos gradientes de intensidade nas imagens. A aplicação da função HOG permite identificar padrões texturais e características de bordas que são essenciais para a detecção de detalhes relevantes nas folhas, independentemente do estado em que se encontram.

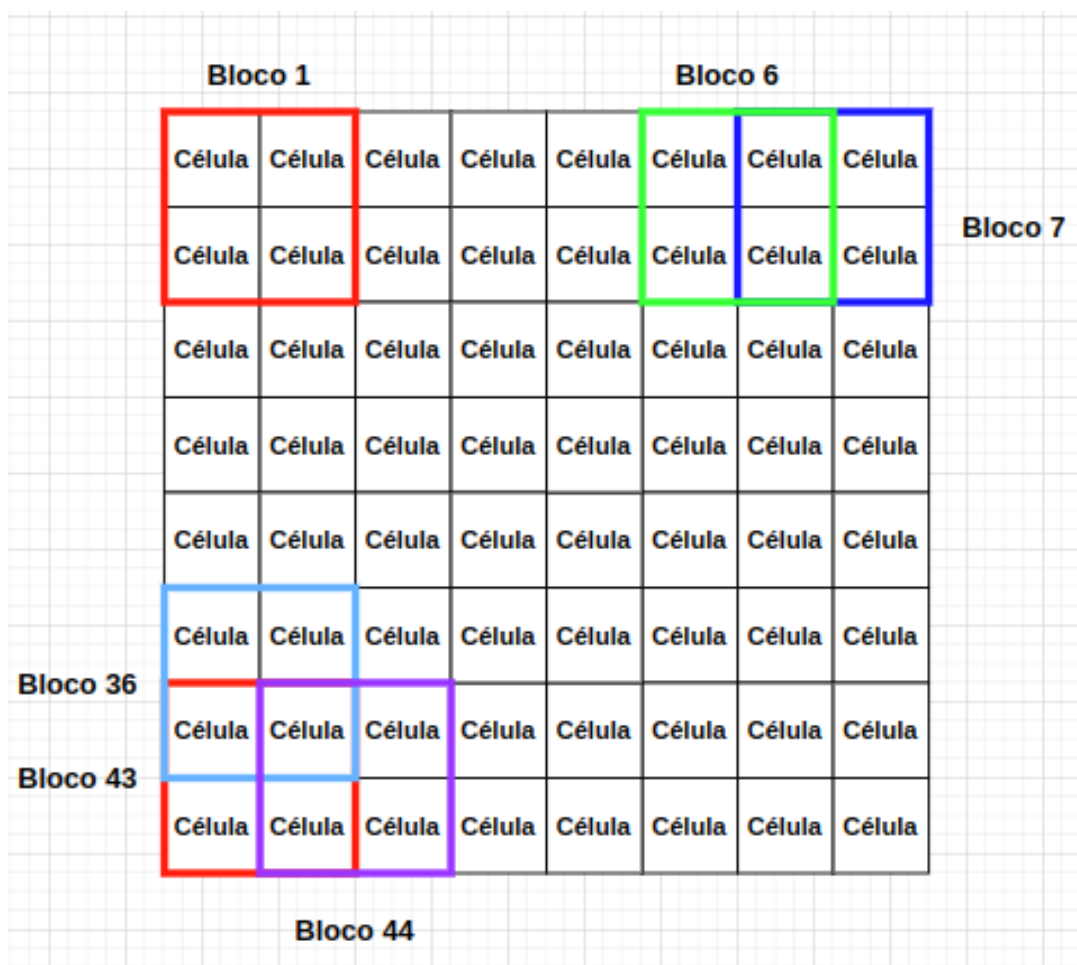


A metodologia HOG opera dividindo a imagem em células menores e, dentro de cada célula, calcula a magnitude e a orientação dos gradientes de intensidade. Posteriormente, essas informações são agrupadas em blocos, formando histogramas de orientações que representam a distribuição das direções dos gradientes em regiões específicas da imagem. A fusão dos histogramas de diferentes blocos resulta em um vetor de características HOG, que captura os elementos fundamentais dos padrões de textura e borda presentes na imagem.

O vetor de características resultante da técnica HOG possui 1764 elementos devido a várias escolhas de configuração. O tamanho das células foi fixado em 8x8 pixels, o que resultou em um total de 64 células na imagem (8 células na horizontal x 8 células na vertical), abrangendo uma área de 64 pixels x 64 pixels. Além disso, a organização dos blocos foi estabelecida como sendo composta por 2x2 células, totalizando assim 49 blocos no conjunto (7 blocos na horizontal x 7 blocos na vertical).

Na Figura 12, a configuração do HOG aplicada a uma imagem. Uma característica importante do HOG é a sobreposição dos blocos, exemplificada nos blocos 6, 7, 36 e 43, permitindo transições suaves entre as regiões da imagem e a captura eficaz de informações locais de textura e forma. Essa sobreposição ajuda a melhorar a representação das características de gradientes ao longo da imagem, tornando o HOG uma técnica eficaz na detecção de objetos em imagens.

Figura 12 – Exemplificação do Blocos e Células



Fonte: O Autor

Com a escolha de 9 *bins* de orientações por célula, cada célula contribui com um vetor de características de 9 elementos, o que é o padrão comumente utilizado na função HOG. Como há 4 células por bloco, isso resulta em um número total de 36 elementos por bloco. Essa configuração é amplamente adotada devido a ser o valor padrão na função HOG, proporcionando uma representação eficaz das orientações dos gradientes em uma imagem e contribuindo para a detecção de características de textura e forma.

Ao multiplicar o número de elementos por bloco (36) pelo número total de blocos (49), obtemos o valor final de 1764 elementos no vetor de características. Portanto, essa combinação específica de células, blocos e *bins* resulta em um vetor de características com 1764 elementos. Esse tamanho resulta da aplicação da função HOG a cada célula da imagem e da subsequente concatenação dos

histogramas de gradientes orientados de cada bloco, formando assim o vetor de características completo.

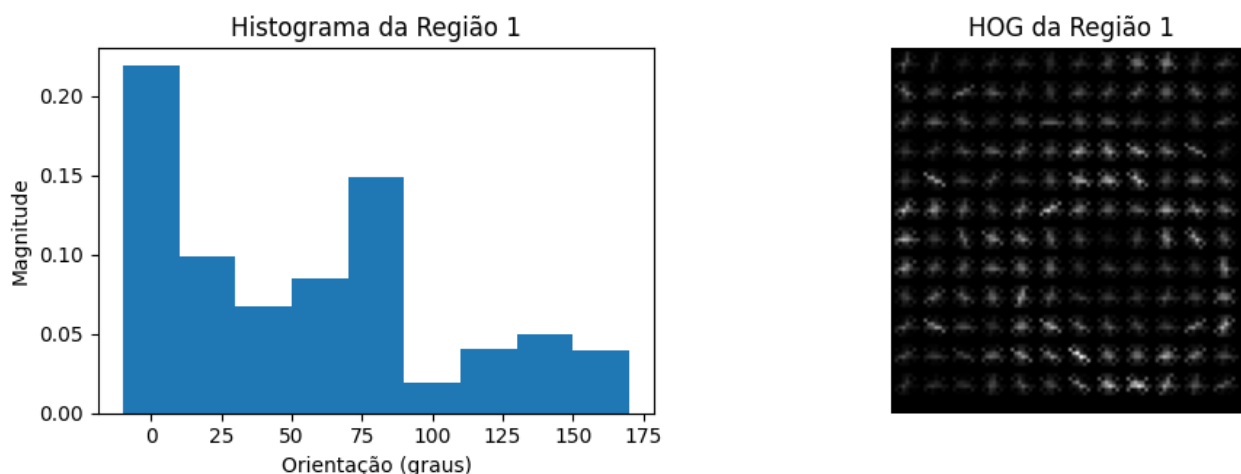
É importante destacar que a técnica HOG oferece uma representação compacta e informativa das características texturais das folhas, permitindo que a rede neural aprenda a distinguir entre diferentes estados de saúde das plantas com base nessas informações. A Figura 13 mostra um exemplo de uma folha com a doença Carijó, com duas regiões destacadas. Após a aplicação da técnica HOG, os resultados obtidos para essas regiões são ilustrados nas Figuras 14 e 15. Mesmo que as duas regiões pertençam à mesma folha, os histogramas obtidos revelam suas diferenças características.

Figura 13 – Imagem de uma folha com doença Carijó com duas regiões destacadas



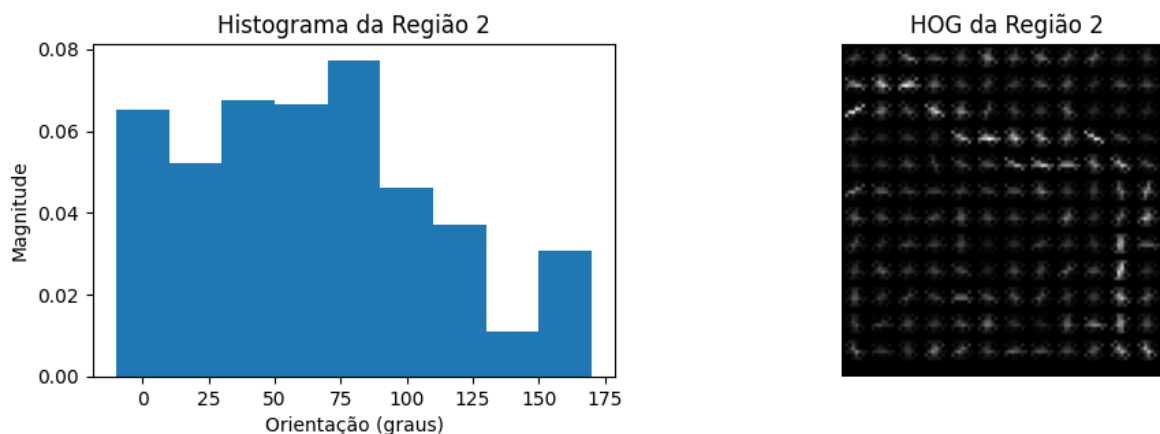
Fonte: o Autor

Figura 14 – HOG obtido para a região 1 da Figura 12



Fonte: o Autor

Figura 15 – HOG obtido para a região 2 da Figura 12



Fonte: o Autor

Além da extração de características através da técnica HOG, também foram utilizadas abordagens de histogramas de cores para complementar a análise das texturas. A função `calcHist` do OpenCV foi empregada para obter histogramas das componentes HSV e RGB das imagens. Esses histogramas representam a distribuição das intensidades de cores em intervalos específicos.

Foram empregados 256 valores para cada canal de cor, resultando em um total de 1536 valores para ambos os espaços de cores. Essa escolha foi feita com o intuito de aumentar a discretização, permitindo uma representação mais precisa das

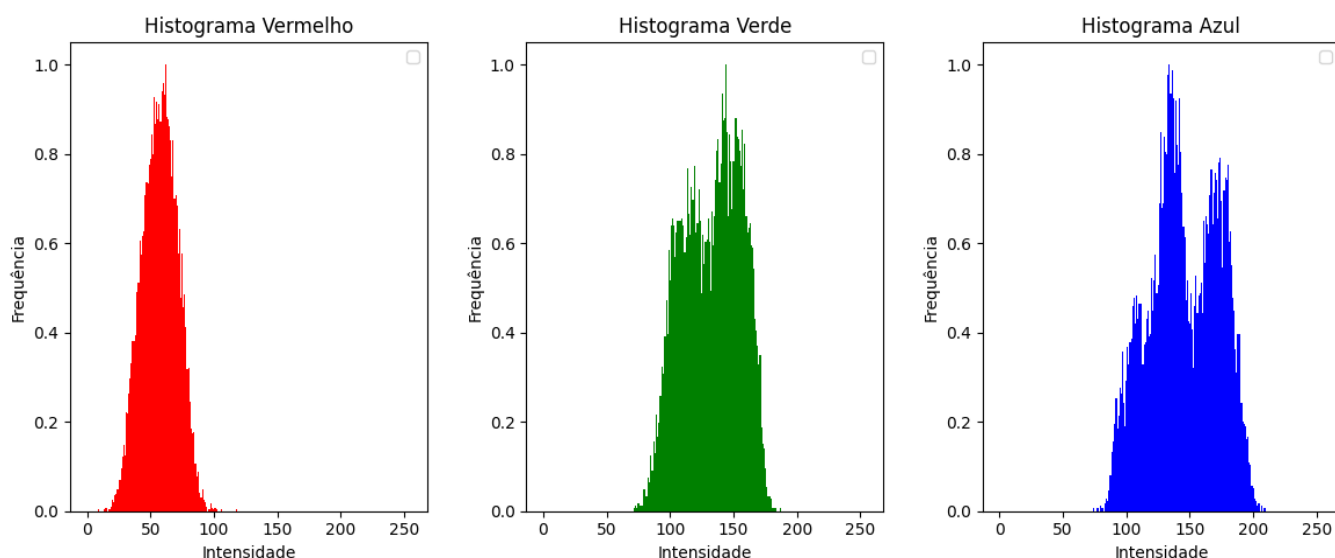
variações de cor na imagem.

Nas Figuras 16 e 17, são observadas diferenças notáveis nos canais RGB entre as regiões 1 e 2 da Figura 13, com destaque para as variações acentuadas nas componentes verde e azul. Essa análise demonstra que duas áreas distintas de uma mesma folha apresentam diferenças significativas nas intensidades dos canais de cores verde e azul.

Nas Figuras 18 e 19, podem ser vistos os canais HSV, com notáveis diferenças nos canais de matiz e valor. Isso evidencia variações substanciais nas características de cor e brilho entre as áreas representadas.

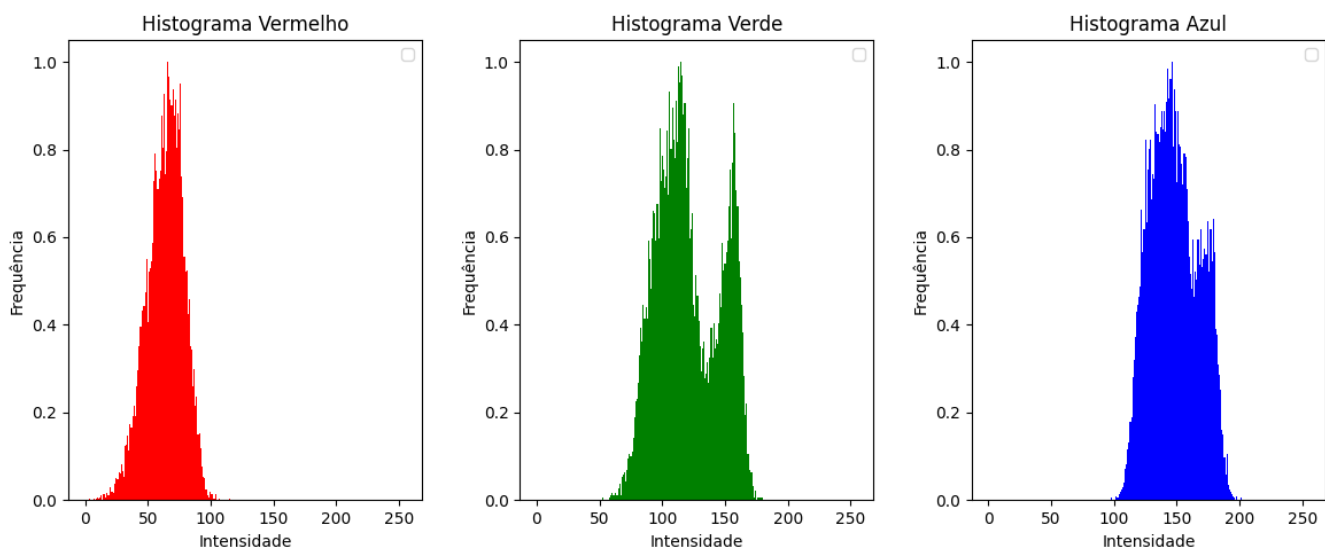
A incorporação desses histogramas de cores oferece uma visão mais abrangente das características das imagens, permitindo que a rede neural capture informações tanto sobre as texturas quanto sobre as distribuições de cores presentes nas folhas, contribuindo assim para uma classificação mais precisa.

Figura 16 – Histogramas RGB obtido para a região 1 da Figura 12



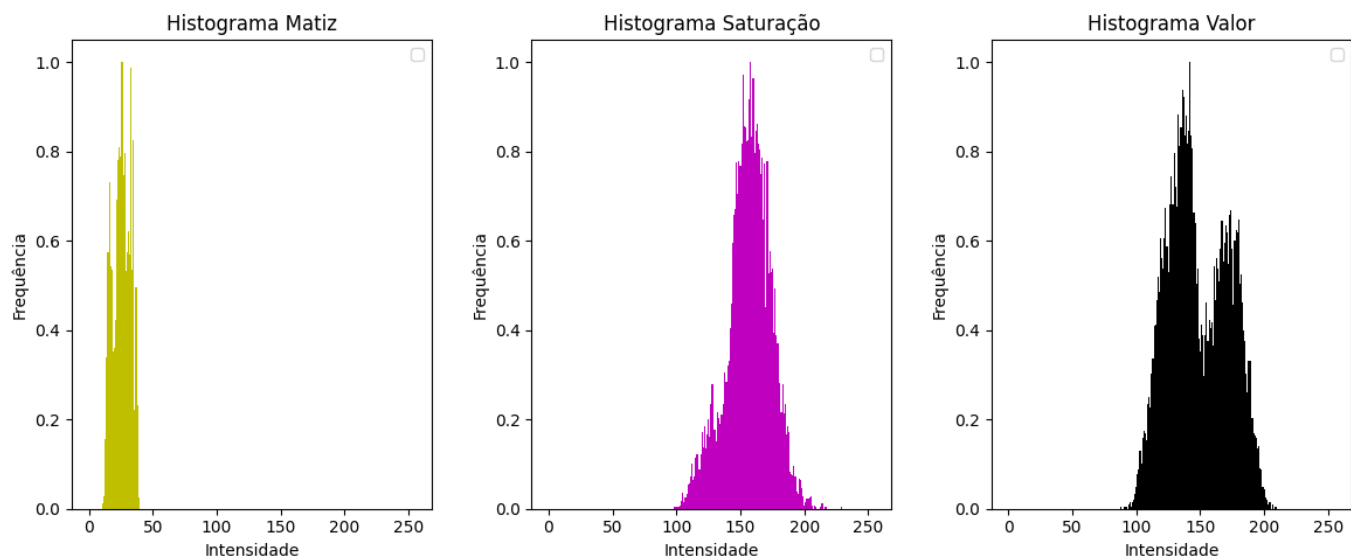
Fonte: o Autor

Figura 17 – Histogramas RGB obtido para a região 2 da Figura 12



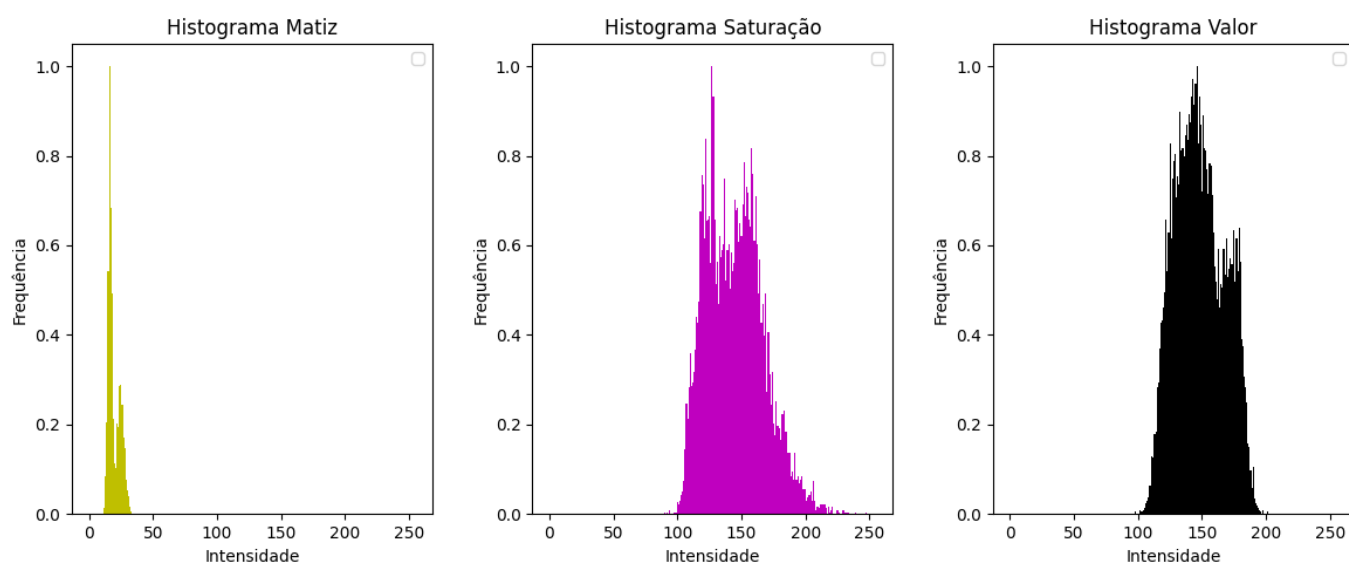
Fonte: o Autor

Figura 18 – Histograma HSV obtido para a região 1 da Figura 12



Fonte: o Autor

Figura 19 – Histograma HSV obtido para a região 2 da Figura 12



Fonte: o Autor

Um passo crucial na extração de características foi a normalização dos valores resultantes dos histogramas. Através da normalização, todos os valores são ajustados para uma mesma escala, tornando-os comparáveis e reduzindo a influência de variações na iluminação ou exposição das imagens. Isso aumenta a robustez das características extraídas e as torna menos sensíveis a flutuações na intensidade de luz. A normalização dos histogramas de cores foi realizada utilizando o método *normalize* do OpenCV, garantindo que as informações extraídas permaneçam coerentes e proporcionem uma representação fiel e discriminativa das texturas presentes nas folhas.

A integração das características obtidas ocorre por meio da combinação de todos os valores em um único vetor. Esse vetor unificado proporciona à rede neural uma representação abrangente e discriminativa das texturas das folhas, ao incluir as características da técnica HOG, bem como os histogramas normalizados das componentes HSV e RGB. Essa fusão de informações texturais e cromáticas capacita a rede neural a diferenciar de maneira eficaz entre folhas saudáveis e doentes, desempenhando um papel fundamental na melhoria da precisão geral do processo de classificação.

### 3.4 TREINAMENTO DA REDE NEURAL

O projeto da rede neural é moldada desde o início do processo, considerando tanto a camada de entrada quanto a camada de saída. Composta por 3300 neurônios, a camada de entrada reflete as características extraídas das imagens processadas, incorporando informações como histogramas de cores e HOG. Essas características são cruciais para a tomada de decisões da rede.

Na camada de saída, 6 neurônios são dispostos, cada um representando uma classe específica de saída. Esses neurônios finais têm a responsabilidade de gerar as probabilidades associadas a cada classe, permitindo que a rede faça previsões sobre a categoria à qual uma determinada entrada pertence. Durante o processo de treinamento, os pesos das conexões entre as camadas são ajustados visando a otimização da capacidade de classificação e a correlação precisa entre as características de entrada e as saídas correspondentes.

Nesse contexto, optou-se por empregar uma abordagem conhecida como Multilayer Perceptron (MLP), um tipo de rede neural artificial. A MLP é caracterizada por suas camadas de neurônios, incluindo uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. A escolha da MLP é justificada por sua comprovada eficácia em lidar com problemas complexos de classificação, como a detecção de doenças em folhas de plantas.

A escolha da MLP é embasada em suas características intrínsecas que a tornam adequada para essa tarefa. A capacidade da MLP de aprender padrões complexos, sua flexibilidade na configuração da arquitetura, sua capacidade de generalização a partir de exemplos de treinamento e sua eficiência no treinamento são fatores decisivos. Além disso, a disponibilidade de bibliotecas de aprendizado de máquina que oferecem a MLP como uma opção prática para implementação e experimentação também é vantajosa.

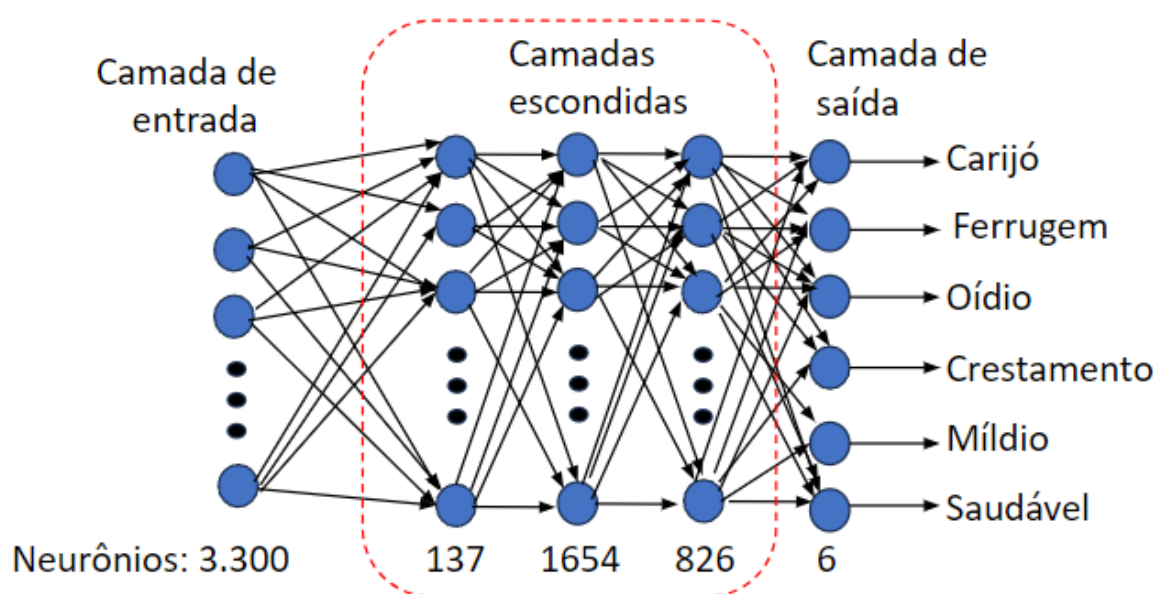
No decorrer do processo de configuração, a técnica de pesquisa em grade (*Grid Search*) é empregada para aprimorar a arquitetura da rede neural. Essa abordagem analisa uma variedade de combinações de hiperparâmetros, avaliando a performance do modelo em diferentes configurações. Os hiperparâmetros são as configurações que controlam o processo de treinamento, por exemplo, a determinação do número de camadas escondidas ou a quantidade de neurônios em cada camada. Através do *Grid Search*, é possível identificar os melhores conjuntos



de hiperparâmetros, que são então utilizados para o refinamento final da rede. Na Tabela 3, apresentada no capítulo Resultados, é exibida a classificação das arquiteturas testadas com base na acurácia da rede em classificar corretamente os conjuntos de imagens. Essa avaliação é crucial para selecionar a combinação de hiperparâmetros que resulte em um desempenho ótimo da rede neural.

Essa estratégia culmina na definição de uma rede neural multicamadas (MLP) personalizada. Com camadas ocultas cujos tamanhos foram otimizados via *Grid Search*, a rede é composta por 137 neurônios na primeira camada, 1654 neurônios na segunda camada e 826 neurônios na terceira camada oculta. Essa configuração foi escolhida após experimentar várias combinações e avaliar os resultados para alcançar um desempenho ideal. Na Figura 20 a representação da rede completa é mostrada.

Figura 20 – Representação esquemática da rede neural escolhida



Fonte: o Autor

Após a fase de treinamento, o modelo da rede neural é armazenado em um arquivo, pronto para uso futuro. Essa abordagem permite a reutilização do modelo sem a necessidade de reexecutar o processo de treinamento, economizando recursos computacionais e tempo.

Com a conclusão bem-sucedida das etapas de treinamento, otimização e armazenamento, o modelo resultante está apto a realizar previsões precisas sobre o estado das folhas. A detecção eficaz de doenças e o apoio à saúde das plantas tornam-se contribuições valiosas proporcionadas por essa rede neural personalizada, especialmente concebida para tarefas de classificação na agricultura.

### 3.5 SEGMENTAÇÃO E PREDIÇÃO

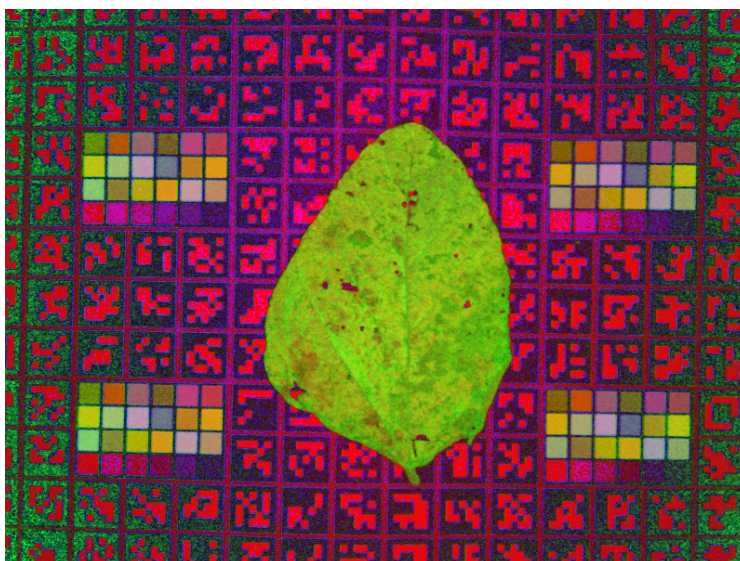
Uma etapa fundamental no processo de detecção de doenças em folhas envolve a segmentação da imagem para isolar a região de interesse, ou seja, a própria folha. Nesse contexto, foi adotada uma abordagem utilizando o canal S do espaço de cores HSV. O canal S é particularmente sensível às variações de cor e saturação, o que o torna uma escolha adequada para a segmentação. A Figura 21 exibe um exemplar de folha com a doença Ferrugem seguida do mesmo exemplar no sistema de cores HSV na Figura 22. Essa mesma folha será utilizada para demonstrar os passos subsequentes.

Figura 21 – Exemplar de folha com a doença Ferrugem



Fonte: EMBRAPA

Figura 22 – Folha com Ferrugem no sistema de cores HSV

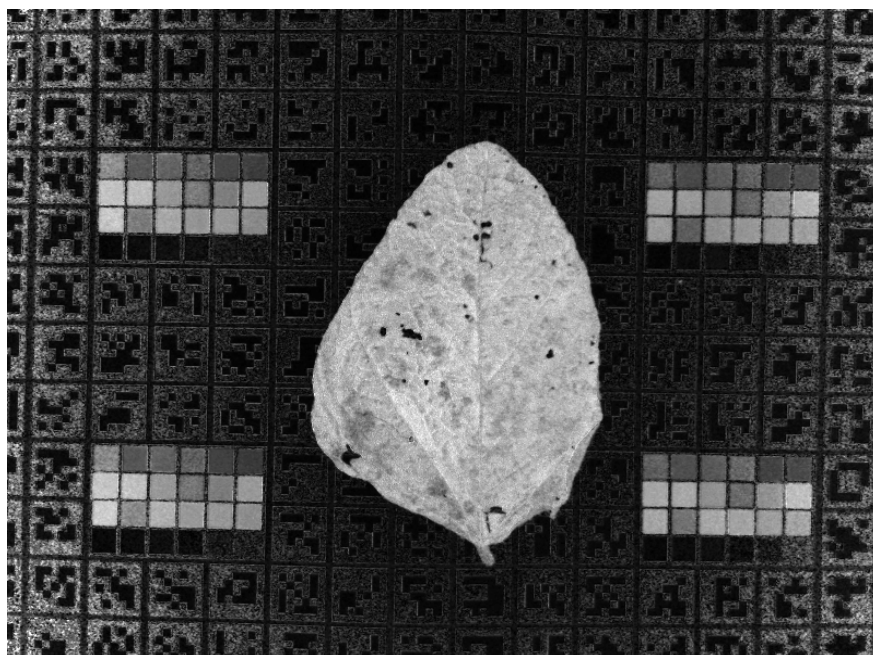


Fonte: O Autor

Para efetuar a segmentação, testes foram conduzidos para estabelecer o limiar mais adequado à separação de imagens de folhas em relação ao fundo, usando o canal S das imagens, que representa a saturação da cor. Após avaliar múltiplos valores, o limiar selecionado foi 0.2, devido à sua eficácia na maioria dos casos ao enfatizar as folhas e mitigar desafios relacionados a variações de iluminação e ruído. O procedimento seguiu uma abordagem de tentativa e erro, resultando na adoção do limiar 0.2 como solução empírica confiável para separar as imagens das folhas do plano de fundo.

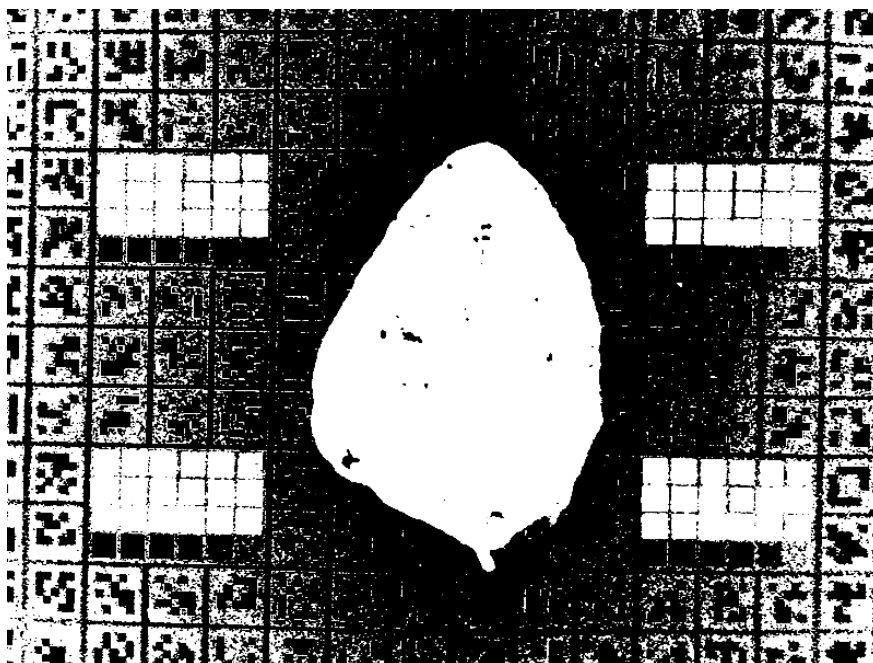
A Figura 23 ilustra um exemplo da influência desse limiar no canal S, destacando a folha em relação ao fundo. Essa seleção do limiar possibilitou a discriminação entre folha e fundo, gerando uma máscara binária que realça a folha e elimina interferências de fundo. A máscara binária foi criada, atribuindo valores acima do limiar à folha e valores abaixo do limiar ao fundo, como exemplificado na Figura 24.

Figura 23 – Folha com Ferrugem na cor S



Fonte: o Autor

Figura 24 – Máscara obtida após filtragem



Fonte: o Autor

Após a obtenção da máscara de segmentação, aplicou-se o método *connectedComponentsWithStats* do OpenCV para identificar regiões contíguas na

máscara. Uma característica importante desse método é que ele não apenas identifica as regiões conectadas, mas também retorna informações sobre cada região, como a área. Essa abordagem permitiu a seleção da maior região, a qual está diretamente associada à imagem da folha, como exemplificado na Figura 25. Durante esse processo, a área de cada região foi utilizada para determinar a sua extensão. Conseqüentemente, a região com a maior área foi escolhida como a região de interesse. Essa região em destaque é apontada como a área onde a análise de texturas e características específicas é realizada, direcionada para a possível detecção de doenças ou anormalidades.

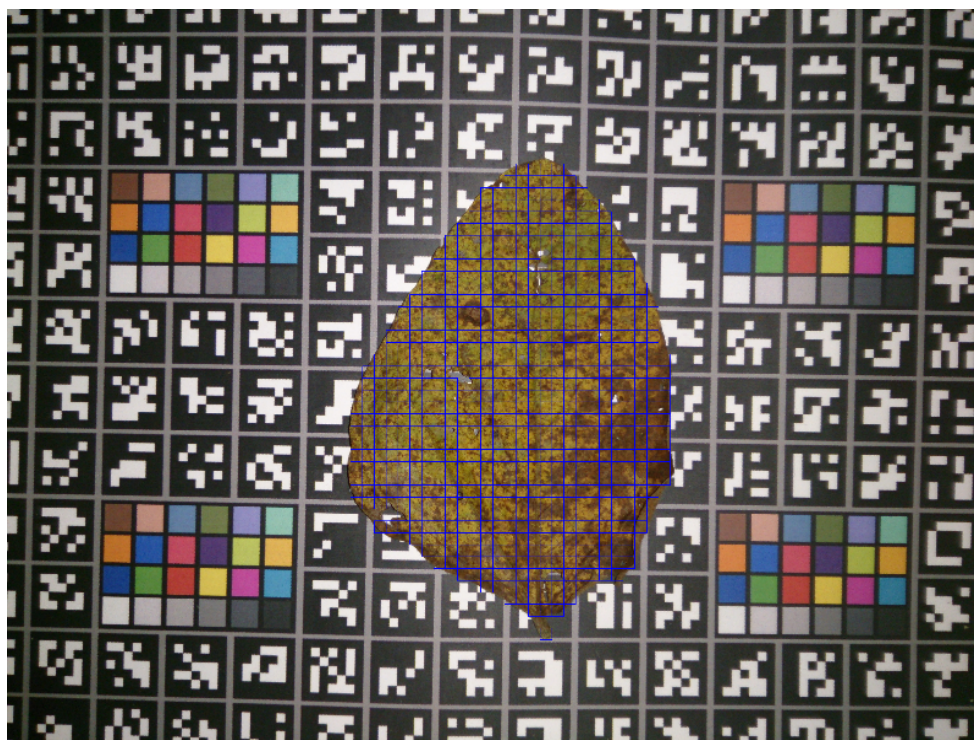
Figura 25 – Imagem obtida após aplicação da máscara



Fonte: o Autor

A partir da imagem isolada da folha, foram extraídas as mesmas características texturais e cromáticas mencionadas anteriormente. No entanto, optou-se por analisar sub-regiões da imagem em vez da imagem completa. Essas sub-regiões foram definidas como quadrados de 64x64 pixels, abrangendo toda a folha. Essa padronização foi realizada para coincidir com o tamanho dos frames de treinamento, garantindo uma análise consistente e correspondente às características utilizadas durante o processo. Pode-se observar na Figura 26 marcados em azul os quadrados separados para análise.

Figura 26 – Imagem original com os quadrados que foram separados para análise

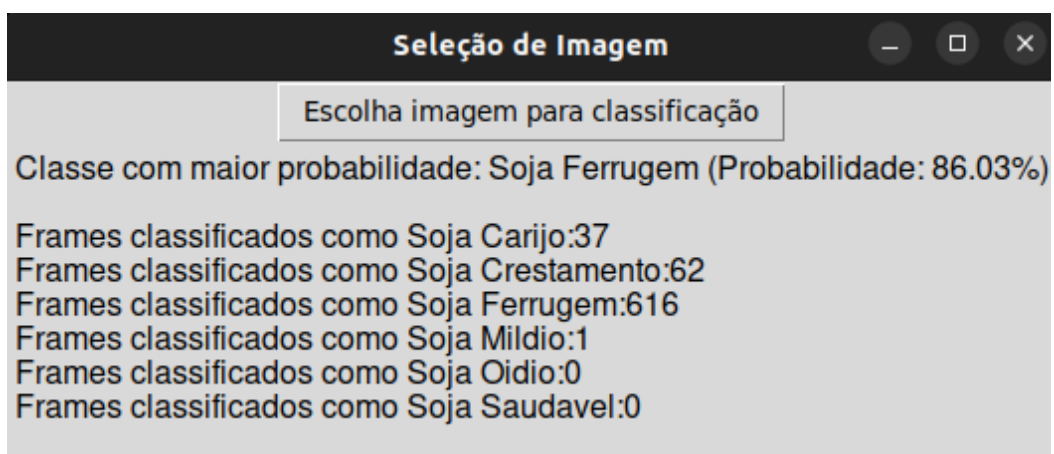


Fonte: O Autor

Primeiro foram descartados os quadrados com mais de 80% de fundo, para os restantes as características foram extraídas e utilizadas para a previsão do estado da folha. A rede neural treinada e otimizada previu a classe de cada sub-região, indicando se a folha naquela área específica está saudável ou doente. Essa abordagem de análise em nível de sub-região permite uma detecção mais detalhada e localizada de possíveis doenças, oferecendo informações valiosas sobre a distribuição e extensão das anomalias na folha.

Concluindo a análise da folha proposta, o resultado exibido na Figura 27 demonstra que 616 quadrados (86,03%) foram categorizados como representativos de casos de Ferrugem.

Figura 27 – Resultados das predições de cada quadrado



Fonte: O Autor

## 4 RESULTADOS

No decorrer deste estudo, foi explorada a aplicação de uma abordagem de um classificador com redes neurais, para a tarefa desafiadora de detecção e classificação de doenças em plantas de soja com base em imagens. A abordagem buscou capturar tanto as características intrínsecas das imagens quanto a incerteza inerente às decisões de classificação. A Tabela 2 mostra o total de imagens usadas para o teste.

Tabela 2 – Número de texturas utilizadas por classe para teste

Classe da Folha	Número de <i>Frames</i>
Carijó	25
Ferrugem	60
Oídio	20
Crestamento	48
Míldio	22
Saudável	18

Fonte: O Autor

Serão abordadas diversas métricas de desempenho que desempenharam um papel crucial na avaliação da eficácia da abordagem adotada para a detecção e classificação de doenças em imagens de plantas de soja. Essas métricas fornecerão uma análise minuciosa do desempenho do modelo, abrangendo diferentes facetas da tarefa de classificação. As métricas selecionadas são as seguintes:

### 4.1 ACURÁCIA

Uma métrica amplamente reconhecida e utilizada para a avaliação de modelos de classificação é a acurácia. Ela quantifica a proporção de previsões corretas em relação ao total de previsões realizadas pelo modelo. No contexto da aplicação em questão, a acurácia proporciona uma visão global da capacidade do modelo de classificar com precisão amostras de plantas saudáveis e doentes.



Contudo, é relevante destacar que a acurácia pode não ser completamente abrangente para avaliar situações de desequilíbrio entre as classes ou quando os erros de classificação têm implicações diferentes.

Para compreender o cálculo da acurácia, é importante entender os seguintes termos:

- VP (Verdadeiros Positivos): Isso se refere aos casos em que o modelo classificou corretamente amostras que pertencem à classe positiva. Em outras palavras, são os casos em que o modelo acertou ao identificar algo como "positivo" e essa identificação estava correta.
- VN (Verdadeiros Negativos): Isso representa os casos em que o modelo classificou corretamente amostras que pertencem à classe negativa. Em outras palavras, são os casos em que o modelo acertou ao identificar algo como "negativo" e essa identificação estava correta.
- FP (Falsos Positivos): Isso ocorre quando o modelo classifica erroneamente amostras da classe negativa como pertencentes à classe positiva. Ou seja, o modelo erra ao identificar algo como "positivo" quando, na verdade, é "negativo".
- FN (Falsos Negativos): Os falsos negativos acontecem quando o modelo classifica erroneamente amostras da classe positiva como pertencentes à classe negativa. Ou seja, o modelo erra ao identificar algo como "negativo" quando, na verdade, é "positivo".

Assim a fórmula da acurácia é dada por:

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

Além de sua utilização na avaliação do desempenho do modelo, a acurácia também desempenhou um papel crucial no processo de escolha da arquitetura da rede neural. Durante a fase de desenvolvimento do modelo, diferentes arquiteturas foram exploradas e treinadas para a tarefa de detecção e classificação de doenças em imagens de plantas de soja. A acurácia foi empregada como uma métrica

orientadora para selecionar a arquitetura mais promissora entre as alternativas avaliadas. Essa abordagem permitiu a identificação da arquitetura que demonstrou maior capacidade de capturar padrões relevantes nas imagens e de generalizar eficazmente para novos dados. Na Tabela 3 apresenta as acurácias obtidas para as diferentes arquiteturas.

Tabela 3 – Acurácia de diferentes arquiteturas da camada oculta

<b>Neurônios Camada Oculta</b>	<b>Acurácia</b>
137, 1654, 826	0,87548
137, 826, 137	0,87044
200, 100, 50	0,86528
413, 826, 413	0,86523
413, 1654, 826	0,86523
1654	0,86517
137, 826	0,86002
275, 826	0,85997
137, 1654	0,85497
550	0,85481
500, 250, 100	0,85476
551	0,84966

Fonte: o Autor

Com base na Tabela 3 a arquitetura escolhida foi de três camadas ocultas com 137, 1654 e 826 neurônios respectivamente, as próximas métricas serão para a arquitetura escolhida.

#### 4.2 PRECISÃO

A equação 2 apresenta a fórmula da precisão, que mede a proporção de previsões positivas corretas (verdadeiros positivos, VP) em relação ao total de previsões positivas (verdadeiros positivos mais falsos positivos, VP + FP). A precisão é uma métrica particularmente útil quando se consideram os casos de

falsos positivos, como no contexto da detecção de doenças. Para a abordagem adotada, obteve-se um valor de precisão de 88,71%, o que indica a capacidade do modelo de minimizar os erros de classificação positivos e realça a sua eficácia na identificação precisa das amostras da classe positiva.

$$Precisão = \frac{VP}{VP + FP} \quad (2)$$

#### 4.3 RECALL

A fórmula do *recall* (revocação) é apresentada na equação 3, e ela mede a proporção de previsões positivas corretas (verdadeiros positivos, VP) em relação ao total de amostras que realmente pertencem à classe positiva (VP + falsos negativos, FN). Essa métrica ganha importância quando o foco está em evitar a não detecção de casos positivos, como no contexto da detecção de doenças. Os resultados obtidos revelaram um recall de 88,08%, destacando a eficácia do modelo em identificar de forma eficaz as amostras da classe positiva, minimizando assim a possibilidade de falsos negativos.

$$RECALL = \frac{VP}{VP + FN} \quad (3)$$

#### 4.3 F1-SCORE

O F1-score é uma métrica que combina a precisão e o *recall* para oferecer uma única medida de desempenho de um modelo de classificação. Essa métrica é especialmente útil quando as classes estão desbalanceadas, o que significa que há uma diferença significativa no número de amostras entre as classes, como é comum em problemas de detecção de doenças em plantas, onde as amostras de plantas saudáveis podem superar em número as plantas doentes ou vice-versa.

A fórmula do F1-score é apresentada na equação 4 e é calculada como a média harmônica entre a precisão e o *recall*. A média harmônica é usada em vez da média aritmética porque é mais sensível a desequilíbrios nas taxas de falsos positivos e falsos negativos, o que é comum em tarefas de classificação desbalanceadas. O F1-score fornece uma medida geral do desempenho do modelo,

encontrando um equilíbrio entre a capacidade de prever corretamente as amostras positivas (precisão) e a capacidade de não perder amostras positivas (recall).

$$F1\ Score = \frac{2 * VP}{2 * VP + FP + FN} \quad (4)$$

No caso mencionado, o modelo obteve um F1-score de 87,82%, o que reflete sua habilidade geral em encontrar um equilíbrio entre a precisão e o *recall*, indicando um desempenho sólido na detecção de doenças em plantas, mesmo em situações de desbalanceamento das classes.

#### 4.4 MATRIZ DE CONFUSÃO

A matriz de confusão é uma ferramenta essencial para a análise detalhada do desempenho de um modelo de classificação. Ela fornece uma visão mais profunda das previsões do modelo em relação às classes reais dos dados, permitindo uma avaliação mais completa dos erros e acertos que o modelo realiza.

A matriz de confusão é uma matriz quadrada em que as linhas representam as classes reais e as colunas representam as classes previstas pelo modelo. Cada célula na matriz indica a contagem de amostras que pertencem à classe da linha real e foram previstas como pertencentes à classe da coluna prevista.

A matriz de confusão obtida se apresenta como mostrado na Tabela 4:

Tabela 4 – Matriz confusão obtida.

	Carijó	Crestamento	Ferrugem	Míldio	Oídio	Saudável
Carijó	15	1	9	0	0	0
Crestamento	0	43	0	0	0	5
Ferrugem	4	0	56	0	0	0
Míldio	0	0	0	18	3	1
Oídio	0	0	0	0	20	0
Saudável	0	0	0	0	0	18

Fonte: o Autor

Interpretando a matriz de confusão:

- Na primeira linha, a classe real é a classe Carijó. O modelo previu corretamente 15 amostras dessa classe. No entanto, houve uma confusão em algumas previsões: 1 amostra da classe Carijó foi erroneamente classificada como classe Crestamento, 9 amostras da classe Carijó foram erroneamente classificadas como classe Ferrugem.
- Na segunda linha, a classe real é a classe Crestamento. O modelo previu corretamente 43 amostras dessa classe. Houve 5 amostras da classe Crestamento que foram erroneamente classificadas como classe Saudável.
- Na terceira linha, a classe real é a classe Ferrugem. O modelo previu corretamente 56 amostras dessa classe. 4 amostras foram erroneamente classificadas como classe Carijó.
- Na quarta linha, a classe real é a classe Míldio. O modelo previu corretamente 18 amostras dessa classe. No entanto, houve confusão em algumas previsões: 3 amostras da classe Míldio foram erroneamente classificadas como classe Oídio, 1 amostra da classe Míldio foi erroneamente classificada como classe Saudável.
- Na quinta linha, a classe real é a classe Oídio. O modelo previu corretamente 20 amostras dessa classe. Não houve confusões nas previsões para essa classe.
- Na sexta linha, a classe real é a classe Saudável. O modelo previu corretamente 18 amostras dessa classe. Não houve confusões nas previsões para essa classe.

Essa matriz permite identificar onde o modelo está se saindo bem e onde está cometendo erros. Por exemplo, observa-se que há confusão entre as classes Carijó e Ferrugem. Essas informações podem nos orientar na otimização do modelo

e na identificação de padrões específicos de erros. Com a combinação da matriz de confusão e outras métricas de desempenho, como acurácia, precisão, recall e F1-score, uma visão abrangente da qualidade do modelo em relação à detecção e classificação de doenças em imagens de plantas de soja é adquirida.. Na Tabela 5, as métricas por classes são apresentadas como uma maneira de avaliar o desempenho do modelo de classificação. As métricas incluem precisão, recall e F1-score para cada classe.

Tabela 5 – Métricas por classe

	Precisão	<i>Recall</i>	F1-score
Carijó	79 %	60 %	68 %
Crestamento	98 %	90 %	93 %
Ferrugem	86 %	93 %	90 %
Míldio	100 %	82 %	90 %
Oídio	87 %	100 %	93 %
Saudável	75 %	100 %	86 %

Fonte: o Autor

A análise da matriz de confusão proporcionou informações valiosas sobre os erros de classificação ocorridos no modelo, indicando direções para possíveis melhorias no futuro. A variabilidade na taxa de *recall* entre as diversas classes ressalta a necessidade de refinamentos específicos, especialmente para doenças com menor presença na amostragem. É importante notar que, dentro desse contexto, as classes Saudável e Carijó merecem uma atenção particular. Ambas apresentaram uma precisão menor que 80%, sugerindo a importância de investigar e abordar os desafios específicos associados a essas classes. Ao direcionar recursos para essas áreas, é possível buscar soluções que melhorem a capacidade do modelo de identificar e classificar adequadamente amostras de plantas nessas categorias. Esse foco direcionado tem o potencial de elevar o desempenho global do modelo em termos de precisão e confiabilidade.

#### 4.5 COMPARATIVO COM OUTROS TRABALHOS

Apesar dos problemas identificados na matriz de confusão, os resultados alcançados ainda se destacam como promissores quando comparados com abordagens tradicionais. O desempenho do modelo desenvolvido demonstrou eficácia na resolução da problemática em questão, especialmente quando contrastado com abordagens mais convencionais de processamento de imagens, como evidenciado no estudo anterior conduzido por Barbedo (2015).

A Tabela 6 apresenta a comparação de quatro classes específicas, exibindo as medidas de precisão obtidas tanto pela referência quanto pelo modelo desenvolvido neste estudo. Essa análise comparativa permite visualizar como o desempenho do modelo se compara ao da abordagem de referência em relação à precisão para essas classes específicas. Essa comparação direta entre os resultados evidencia as vantagens ou melhorias proporcionadas pelo modelo proposto em relação às abordagens tradicionais.

No caso do Crestamento, o modelo desenvolvido apresenta um aumento significativo de 35.7% para 98% em termos de precisão. Isso aponta para uma melhoria drástica na capacidade do modelo em identificar e classificar essa classe. No contexto da Ferrugem, também é evidente um avanço considerável, com a precisão aumentando de 71% para 86%. Quando se trata da classe Míldio, a precisão atinge 100% no modelo desenvolvido em comparação com 91.3%. Similarmente, na classe Oídio, o modelo atual apresenta uma precisão de 87%, enquanto o modelo de Barbedo (2015) registrou apenas 61.8%. Isso representa uma melhoria considerável na detecção do Oídio.

Tabela 6 – Comparativo da precisão

Classe da Folha	Precisão (BARBEDO, 2015)	Precisão (Autor)
Crestamento	35.7%	98%
Ferrugem	71%	86%
Míldio	91.3%	100%
Oídio	61.8%	87%

Fonte: o Autor

## 5 CONSIDERAÇÕES FINAIS

### 5.1 CONCLUSÕES

Os objetivos iniciais deste estudo foram plenamente alcançados. A aplicação de técnicas de redes neurais permitiu a criação de um modelo robusto capaz de analisar imagens de folhas de soja e classificar com sucesso diferentes classes de doenças, além de determinar a saúde geral das plantas. A taxa de acurácia alcançada, atingindo 87,54%, é um reflexo sólido da capacidade do modelo em realizar classificações corretas, mesmo em meio à complexidade das classes e à diversidade das imagens.

A avaliação detalhada das métricas de precisão, recall e F1-score trouxe uma visão abrangente do desempenho do modelo em cada classe, evidenciando sua habilidade em identificar padrões e tomar decisões de classificação acertadas. A matriz de confusão revelou áreas específicas de melhoria, principalmente nas classes com menor representação, reforçando a importância de futuras otimizações.

No entanto, é importante ressaltar que há espaço para aprimoramentos e expansões. A coleta de dados mais abrangente e representativa, juntamente com um conjunto de imagens de treinamento mais extenso, têm o potencial de aperfeiçoar ainda mais a precisão do modelo. A análise mais aprofundada dos hiperparâmetros das redes neurais pode resultar em refinamentos significativos.

### 5.2 TRABALHOS FUTUROS

Em relação aos possíveis avanços futuros nesta área de pesquisa, existem diversas oportunidades para aprimorar e expandir ainda mais a detecção de doenças em folhas. Um caminho promissor é a utilização de arquiteturas de redes neurais mais avançadas, como as Convolutional Neural Networks (CNNs), para melhorar o processo de análise e identificação das partes doentes das plantas.

Ao invés de analisar a folha sub-região por sub-região, as CNNs podem ser empregadas para realizar a segmentação e detecção de doenças em nível de pixel. As CNNs são especialmente projetadas para trabalhar com dados de imagem e podem aprender padrões complexos a partir dos dados de treinamento. Essa



abordagem permitiria identificar de forma mais precisa as áreas afetadas pelas doenças, resultando em uma detecção mais refinada e localizada.

Além disso, um aspecto crucial para a melhoria contínua da precisão do modelo é a expansão do banco de dados de imagens. Um banco de dados mais amplo e diversificado, contendo uma variedade de tipos de folhas, diferentes doenças e várias condições de crescimento, é fundamental para garantir que o modelo seja capaz de generalizar bem para situações do mundo real. O aumento do banco de dados proporcionaria à rede neural uma exposição mais abrangente às variações e nuances das folhas, o que poderia levar a uma melhor capacidade de generalização e, conseqüentemente, a previsões mais confiáveis.

## REFERÊNCIAS

ACADEMIA EDU. **Detecção automática de doenças em folhas de plantas baseado em redes de Petri e algoritmos genéticos**. 2016. Disponível em: [https://www.academia.edu/22107570/Detec%C3%A7%C3%A3o\\_autom%C3%A1tica\\_de\\_doen%C3%A7as\\_em\\_folhas\\_de\\_plantas\\_baseado\\_em\\_redes\\_de\\_Petri\\_e\\_algoritmos\\_gen%C3%A9ticos](https://www.academia.edu/22107570/Detec%C3%A7%C3%A3o_autom%C3%A1tica_de_doen%C3%A7as_em_folhas_de_plantas_baseado_em_redes_de_Petri_e_algoritmos_gen%C3%A9ticos). Acesso: 05 ago. 2023

BARBEDO, C. V. G. J. G. A. **Automatic classification of soybean diseases based on digital images of leaf symptoms**. In: X congresso Brasileiro De Agroinformática. [s.n.], 2015.

Disponível em: <<http://www.alice.cnptia.embrapa.br/alice/handle/doc/1028847>>. Acesso em 20 jul. 2023.

CAPES. **Redes Neurais Artificiais: Uma Abordagem Para Sala de Aula**. Disponível em: <https://educapes.capes.gov.br/bitstream/capes/432794/1/e-book-Redes-Neurais-Artificiais-uma-Abordagem-para-Sala-de-Aula.pdf>. Acesso: 05 ago. 2023

CHAUDHARY, P. et al. **Color transform based approach for disease spot detection on plant leaf**. International Journal of Computer Science and Telecommunications, v. 3, n. 6, p. 65–70, 2012.

Confederação da Agricultura e Pecuária do Brasil. **Panorama do Agro**. Disponível em: <<https://www.cnabrazil.org.br/cna/panorama-do-agro>>. Acesso em: 13 jun 2023.

CRM PIPE RUN. **Visão computacional: o que é, aplicações e importância**. Disponível em: <https://crmpiperun.com/blog/visao-computacional/>. Acesso em 24 jul. 2023.

EMBRAPA ALGODÃO. **Guia de identificação e controle das principais doenças do algodoeiro no Estado de Goiás**. Campina Grande: Embrapa Algodão, 2003. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/CNPA/16076/1/DOC113.PDF>. Acesso em 26 jul. 2023.

FERREIRA, M. D.; SPRICIGO, P. C. **Colorimetria - princípios e aplicações na agricultura**. Embrapa Instrumentação, São Carlos, SP, 2017. Disponível em: 6. Acesso em: 6 ago. 2023.

FUENTES, A. et al. **A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition**. Sensors, Switzerland, 2017.

GARRIDO, L. da R.; HOFFMANN, A.; SILVEIRA, S. V. da (Ed.). **Produção integrada de uva para processamento: manejo de pragas e doenças**. Brasília, DF:

Embrapa, 2015. v. 4, cap. 2, p. 33-48. Acesso em 15 jul. 2023.

GONZALEZ, R. C.; WOODS, R. E. **Processamento digital de imagens**. 3. ed. São Paulo: Pearson Prentice Hall, 2010.

HASSAN, M. ul. VGG16. **Convolutional Network for Classification and Detection**. 2018. Disponível em: <<https://neurohive.io/en/popular-networks/vgg16/>>. Acesso 20 jul. 2023.

HAYKIN simon. **Redes Neurais Princípios e Prática**. Porto Alegre: Bookman, 2001.

IBM. **O que são Redes Neurais?** Disponível em: <https://www.ibm.com/br-pt/topics/neural-networks>. Acesso em 03 ago 2023.

MEDIUM. **Redes Neurais Artificiais. Fundamentos teóricos**. Disponível em: <https://medium.com/@avinicius.adorno/redes-neurais-artificiais-5b65a43614a0>. Acesso em 01 ago. 2023.

MICHEREFF,S.J. **Fundamentos de fitopatologia**. Recife: UFRPE , 2001. Disponível em:<https://www.bibliotecaagptea.org.br/agricultura/defesa/livros/FUNDAMENTOS%20DE%20FITOPATOLOGIA.pdf>. Acesso em 02 ago. 2023.

MICROSOFT. **RGB Color Spaces**. Disponível em: <https://learn.microsoft.com/en-us/windows/win32/wcs/rgb-color-spaces>. Acesso em 05 ago. 2023.

MICROSOFT. **Color**. Disponível em: <https://learn.microsoft.com/pt-br/windows/win32/uxguide/vis-color>. Acesso em 05 ago. 2023.

NVIDIA. **O que é Visão Computacional?** Disponível em: <https://www.nvidia.com/pt-br/glossary/data-science/computer-vision/>. Acesso em 05 ago. 2023.

PIRES, R. D. L. et al. **Local descriptors for soybean disease recognition**. Computers and Electronics in Agriculture, v. 125, p. 48–55, 2016. ISSN 0168-1699. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S016816991630179X>. Acesso em 16 jul 2023.

PIXFORCE. **Visão Computacional: Um Guia Completo para tirar suas dúvidas**. Disponível em: <https://www.pixforce.com.br/post/vis%C3%A3o-computacional-um-sistema-de-vis%C3%A3o-inteligente>. Acesso em 16 jul 2023.

PORTAL SÃO FRANCISCO. **Fitopatologia, Doenças, Plantas, O que é Fitopatologia.** Disponível em: <https://www.portalsaofrancisco.com.br/biologia/fitopatologia>. Acesso em 15 jul. 2023.

RAVAL,A.R.; SARTORATO,A. **Conceitos básicos sobre doenças de plantas.** Goiânia: Embrapa-CNPAF, 1994. Disponível em: <https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/199892/1/CNPAF1994p7.pdf>. Acesso 25 jul. 2023.

SAS. **Visão computacional: o que é e qual sua importância?** Disponível em: [https://www.sas.com/pt\\_br/insights/analytics/computer-vision.html](https://www.sas.com/pt_br/insights/analytics/computer-vision.html). Acesso em 04 ago. 2023.

Singh, A. **Feature Engineering for Images: A Valuable Introduction to the HOG Feature Descriptor**

<https://www.analyticsvidhya.com/blog/2019/09/feature-engineering-images-introduction-hog-feature-descriptor/>. Acesso 22 ago 2023

USP. **Deep learning em dois estágios para detecção e classificação de doenças em folhas de plantas.** 2022. Disponível em: <https://www.teses.usp.br/teses/disponiveis/55/55134/tde-02022022-121135/?&lang=pt-br>. Acesso 02 ago. 2023.

