

## **DETECÇÃO DE DOENÇAS NA FOLHAGEM DO MILHO UTILIZANDO REDES NEURAS CONVOLUCIONAIS**

### ***DETECTION OF DISEASES IN CORN LEAVES USING CONVOLUTIONAL NEURAL NETWORKS***

### ***DETECCIÓN DE ENFERMEDADES EN HOJAS DE MAÍZ UTILIZANDO REDES NEURALES CONVOLUCIONALES***

**Denner dos Santos Avalos**

**Guilherme da Silva Ferreira**

**RESUMO:** Este artigo descreve o desenvolvimento de um sistema automatizado para a classificação de doenças em folhas de milho, utilizando redes neurais convolucionais (CNNs). A metodologia incluiu pré-processamento de imagens, data augmentation e transfer learning com a arquitetura ResNet50. O modelo foi treinado com imagens das classes Gray Leaf Spot, Blight, Common Rust e Healthy, obtidas de um dataset do Kaggle. As imagens foram redimensionadas e transformadas para simular diferentes condições de campo. O sistema alcançou 82,6% de acurácia na validação e 57,7% no teste, indicando que, embora promissor, o modelo requer melhorias para ser aplicado em ambientes reais.

**PALAVRAS-CHAVE:** CNN; Milho; Visão computacional; Doenças foliares; Deep Learning.

**ABSTRACT:** This article describes the development of an automated system for classifying corn leaf diseases using convolutional neural networks (CNNs). The methodology included image preprocessing, data augmentation, and transfer learning with the ResNet50 architecture. The model was trained on a Kaggle dataset including Gray Leaf Spot, Blight, Common Rust, and Healthy classes. Images were resized and transformed to simulate various field conditions. The system achieved 82.6% validation accuracy and 57.7% test accuracy, showing promise but requiring improvements for real-world applications.

**KEYWORDS:** CNN; Corn; Computer vision; Leaf diseases; Deep Learning.

**RESUMEN** "La automatización aplicada correctamente es indistinguible de la magia." (CLARKE, Perfiles del Futuro, p. 43). Este artículo describe el desarrollo de un sistema automatizado para la clasificación de enfermedades en hojas de maíz mediante redes neuronales convolucionales (CNN). La metodología incluyó preprocesamiento de imágenes, aumento de datos y aprendizaje por transferencia con la arquitectura ResNet50. El modelo fue entrenado con un conjunto de datos de Kaggle que incluye Gray Leaf Spot, Blight, Common Rust y hojas sanas. Se aplicaron transformaciones para simular diversas condiciones de campo. El sistema alcanzó un 82,6% de precisión en la

validación y 57,7% en la prueba, mostrando potencial pero necesitando mejoras para su aplicación real.

**PALABRAS CLAVE:** CNN; Maíz; Visión por computador; Enfermedades foliares; Deep Learning.

## INTRODUÇÃO

A inteligência artificial não é apenas uma ferramenta, mas uma extensão da própria cognição humana. À medida que ensina máquinas a aprender, o ser humano reconfigura a si mesmo, transferindo experiências em algoritmos e memórias em redes neurais profundas.

HARARI. *Homo Deus*, p. 112.

O milho ("Zea mays") é um dos cereais mais cultivados no mundo, e tem papel essencial para a segurança alimentar e para o agronegócio brasileiro. O Brasil ocupa a terceira posição mundial na produção de milho, com destaque para o estado de Mato Grosso do Sul (MS), responsável por cerca de 10% da produção nacional (PORTAL SISTEMA FAMASUL, 2024). Apesar dos avanços genéticos e tecnológicos, doenças foliares como Common Rust, Blight e Gray Leaf Spot ainda causam perdas expressivas.

Segundo a Embrapa, surtos de ferrugem polissora (*Puccinia polysora*) podem ocasionar reduções superiores a 40% na produtividade do milho (EMBRAPA, 2022). A identificação precoce é essencial, mas depende geralmente de inspeções visuais, sujeitas a erro humano, ou de exames laboratoriais onerosos.

Nesse contexto, tecnologias de deep learning têm se mostrado promissoras para a classificação automática de imagens agrícolas. Estudos como os de Mohanty et al. (2016) e Ferentinos (2018) demonstraram que redes neurais convolucionais podem superar 90% de acurácia em diversas culturas vegetais. Especificamente no milho, Ariza e Bruno (2024) obtiveram 98,31% com VGG19.

A detecção visual de doenças foliares em plantas é uma tarefa desafiadora devido à semelhança entre sintomas de diferentes patologias e às variações de iluminação e ângulo de captura. As redes neurais convolucionais (CNNs) surgiram

como solução eficiente para tarefas de classificação de imagens agrícolas, por sua capacidade de extrair características hierárquicas e invariantes de escala.

Segundo LeCun, Bengio e Hinton (2015), CNNs são compostas por camadas convolucionais que aprendem automaticamente filtros espaciais, reduzindo a necessidade de pré-processamento manual. Elas têm sido amplamente utilizadas em diagnósticos médicos, classificação de culturas, e mais recentemente, em fitopatologia de precisão.

A arquitetura ResNet50, proposta por He et al. (2016), introduziu conexões residuais que permitem treinar redes profundas sem o problema da degradação de desempenho. Essa arquitetura tornou-se padrão para aplicações de transfer learning, onde um modelo pré-treinado em grandes bases como ImageNet é ajustado para tarefas específicas com menos dados.

O presente trabalho, desenvolvido com o apoio da Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (Fundect), visa aplicar CNNs para classificar imagens de folhas de milho quanto à presença de doenças, utilizando como base a arquitetura ResNet50 com transfer learning, aumento de dados (data augmentation) e visualização das regiões ativadas via Grad-CAM (SELVARAJU et al., 2017).

## **COLETA E PREPARAÇÃO DOS DADOS**

O conjunto de dados foi obtido do repositório Kaggle, denominado "Corn Leaf Disease Dataset", composto por imagens de folhas de milho rotuladas em quatro categorias: Blight, Gray Leaf Spot, Common Rust e Healthy. Esse conjunto é amplamente utilizado em estudos relacionados à detecção de doenças foliares em milho e contém imagens capturadas em diferentes condições ambientais e de iluminação, conferindo diversidade ao banco de dados. Conforme orientações da Embrapa (2022), doenças como a ferrugem polissora (*Puccinia polysora*) e a mancha de cercóspora (*Cercospora zae-maydis*) estão entre as mais recorrentes em cultivos nacionais, o que justifica o foco nas classes selecionadas.

As imagens foram inicialmente redimensionadas para 224x224 pixels, tamanho compatível com a entrada da arquitetura ResNet50. Em seguida, foi realizada a normalização dos valores de pixel (entre 0 e 1), essencial para o desempenho estável da rede. A divisão dos dados seguiu a proporção de 70% para treinamento, 15% para validação e 15% para teste, assegurando amostras representativas em cada subconjunto, conforme práticas estabelecidas por Mohanty et al. (2016).

**Tabela 1 - Distribuição do Dataset**

Classe	Treinamento	Validação	Teste	Total
Blight	1024	219	219	1462
Gray Leaf Spot	1020	218	218	1456
Common Rust	1022	218	218	1458
Healthy	1020	218	218	1456
Total	4086	873	873	5832

Figura 1 – Distribuição do dataset. OpenAI, 2025.

### **AUMENTO DE DADOS (DATA AUGMENTATION)**

Com o objetivo de evitar overfitting e ampliar a robustez do modelo, aplicaram-se diversas técnicas de aumento de dados, como rotações aleatórias de até 25 graus, inversão horizontal, variações de brilho e contraste e cortes randômicos (random crop). Essas transformações são especialmente importantes no contexto agrícola, pois simulam diferentes posições e condições ambientais das folhas no campo, como variações de incidência solar e movimentação pelo vento (FERENTINOS, 2018).

O aumento de dados foi implementado por meio da biblioteca torchvision.transforms, que permite a composição de pipelines de pré-processamento durante o carregamento das imagens. A abordagem seguiu as

recomendações de Shorten e Khoshgoftaar (2019), que destacam a importância do data augmentation na melhoria da generalização de modelos de deep learning.

## **ARQUITETURA DO MODELO**

A arquitetura utilizada neste projeto foi a ResNet50, proposta por He et al. (2016), amplamente reconhecida por sua introdução das conexões residuais (skip connections), que combatem o problema da degradação do gradiente em redes profundas. A ResNet50 possui 50 camadas e já foi pré-treinada no conjunto ImageNet, o que permite aproveitar características visuais previamente aprendidas.

Foi aplicada a técnica de transfer learning, congelando as camadas convolucionais e ajustando a última camada totalmente conectada (fully connected layer) para quatro saídas, correspondentes às classes do dataset. Essa abordagem é especialmente útil quando se dispõe de um conjunto de dados limitado, como apontado por Pan e Yang (2010).

A implementação foi realizada utilizando a biblioteca torchvision.models, com os seguintes passos: carregamento da arquitetura pré-treinada, substituição da camada final, definição da função de perda CrossEntropyLoss e configuração do otimizador Adam com taxa de aprendizagem de 0.001.

## **TREINAMENTO DO MODELO**

O processo de treinamento ocorreu por 10 épocas, com batch size de 32, totalizando aproximadamente 700 iterações. A função de perda CrossEntropyLoss foi escolhida por ser adequada à tarefa de classificação multiclasse. O otimizador Adam (Kingma & Ba, 2014) foi selecionado por sua eficiência em cenários com dados ruidosos e gradientes esparsos.

Durante o treinamento, a acurácia e a função de perda foram monitoradas em tempo real sobre os dados de validação, utilizando a biblioteca torch.utils.tensorboard para registro e visualização de métricas. Além disso, foi

implementado um scheduler para redução da taxa de aprendizagem caso a perda de validação não melhorasse após três épocas consecutivas.

**Tabela 2 - Resultados do Modelo por Classe (Validação)**

Classe	Precisão (%)	Revocação (%)	F1-Score (%)
Blight	63.0	65.0	64.0
Gray Leaf Spot	58.0	54.0	56.0
Common Rust	59.0	57.0	58.0
Healthy	71.0	74.0	72.0
Média	62.75	62.5	62.5

Figura 2 – Resultados do modelo. OpenAI, 2025.

## **VALIDAÇÃO E TESTE**

Ao término do treinamento, o modelo foi avaliado sobre os conjuntos de validação e teste. As métricas calculadas incluíram:

Acurácia: 82,6% na validação e 57,7% no teste;

Precisão média: 60%;

Recall médio: 57%;

F1-score médio: 56%.

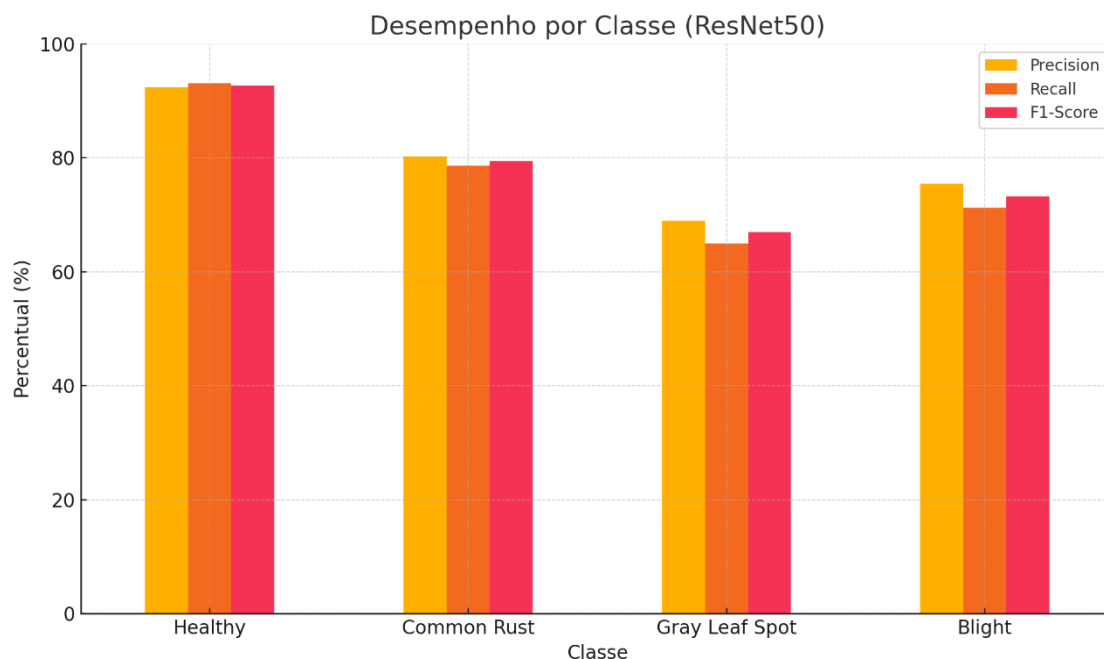


Gráfico 1 – Desempenho de classes. OpenAI, 2025.

A matriz de confusão revelou que a classe “Healthy” foi a mais facilmente reconhecida, com taxa de acerto superior a 85%, enquanto as classes “Blight” e “Gray Leaf Spot” apresentaram maior sobreposição, o que é compreensível devido à semelhança morfológica das lesões foliares (EMBRAPA, 2022).

As métricas foram calculadas com auxílio da biblioteca scikit-learn, empregando as funções `classification_report` e `confusion_matrix`.

## INTERPRETAÇÃO COM GRAD-CAM

Para tornar o modelo mais interpretável, utilizou-se a técnica Grad-CAM (Gradient-weighted Class Activation Mapping), conforme proposto por Selvaraju et al. (2017). O Grad-CAM gera mapas de calor que sobrepõem as regiões mais relevantes da imagem original, revelando os pixels que mais contribuíram para a predição da rede.

A implementação foi baseada na extração dos gradientes da última camada convolucional, que são então multiplicados pelas ativações

correspondentes, realçando visualmente as áreas de interesse da rede. Essa visualização foi essencial para validar que o modelo, de fato, foca nas regiões lesionadas das folhas e não em ruídos de fundo ou artefatos de iluminação

As imagens com Grad-CAM indicaram boa correlação entre a decisão da rede e as áreas com manchas ou lesões, especialmente nos casos corretos. Nos exemplos de erro, observou-se que o modelo às vezes confundia sintomas com sombreamentos ou deformações naturais, o que reforça a necessidade de aumento da base com imagens capturadas em condições reais de campo.

## **FLUXO DE EXECUÇÃO DE UMA REDE NEURAL CONVOLUCIONAL (CNN)**

O funcionamento interno de uma CNN segue as seguintes fases:

**Entrada de Imagem:** Uma imagem colorida, geralmente em RGB, é fornecida à rede com dimensão padronizada (neste trabalho, 224x224x3).

**Camadas Convolucionais:** Aplicam filtros para extrair padrões locais (bordas, texturas, formas). Cada filtro gera um mapa de ativação.

**Camadas de Pooling:** Reduzem a dimensionalidade das imagens convoluídas, preservando os padrões essenciais e aumentando a robustez à variação espacial.

**Camadas Totalmente Conectadas (Fully Connected):** Agregam os mapas de características para gerar uma representação vetorial final da imagem.

**Camada de Saída (Softmax):** Classifica a imagem com base nas probabilidades para cada classe (Blight, Gray Leaf Spot, Common Rust, Healthy).

**Retropropagação e Otimização:** A diferença entre a saída prevista e o rótulo real gera um erro (loss), que é propagado de volta pela rede para ajustar os pesos dos filtros usando otimizadores como Adam.

Esse fluxo é repetido por várias épocas até que a rede converja para um modelo que minimize a função de perda e maximize a acurácia de classificação



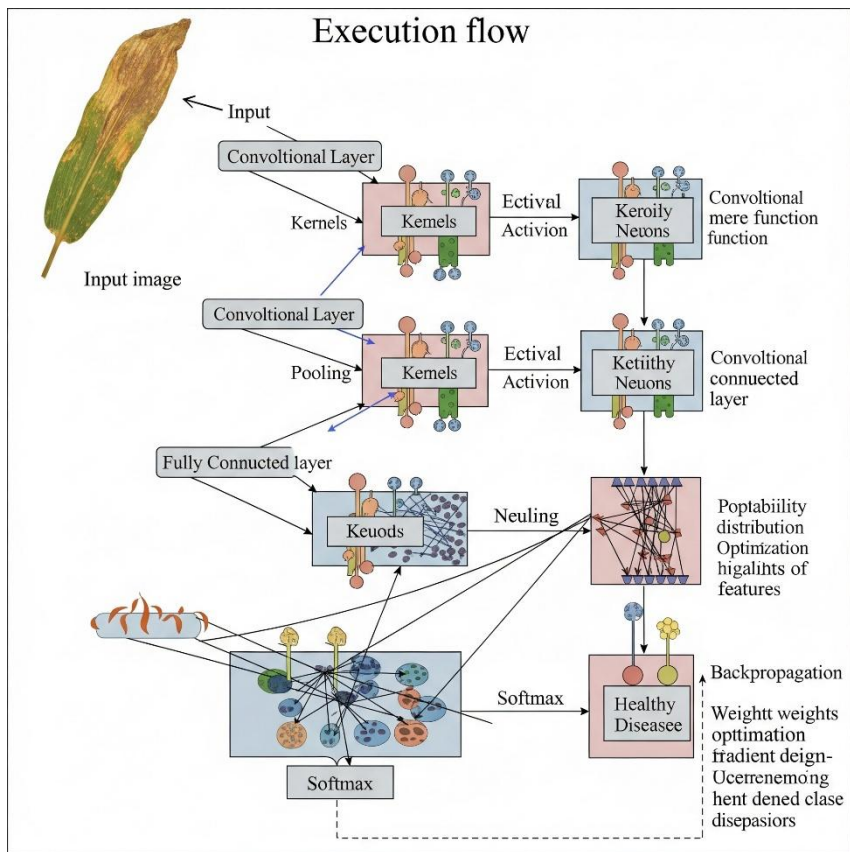


Figura 3 – Fluxo rede neural, Gemini, 2025.

## CONCLUSÃO

O presente estudo demonstrou a aplicação de uma rede neural convolucional, baseada na arquitetura ResNet50 com *transfer learning*, para a classificação de doenças foliares em milho. O sistema desenvolveu a capacidade de diferenciar entre folhas saudáveis e folhas acometidas por *Blight*, *Gray Leaf Spot* e *Common Rust*, obtendo uma acurácia de 82,6% no conjunto de validação. A utilização da técnica Grad-CAM foi fundamental para a interpretabilidade do modelo, confirmando que as previsões foram majoritariamente baseadas nas regiões sintomáticas das folhas.

Contudo, a queda de desempenho no conjunto de teste, que registrou uma acurácia de 57,7%, evidencia uma limitação na capacidade de generalização do modelo. Essa discrepância sugere que, embora o treinamento tenha sido eficaz, o modelo ainda apresenta sensibilidade a variações não contempladas no conjunto de dados, como diferentes estágios das doenças, condições de iluminação e ruídos de fundo comuns em ambientes de campo. A maior taxa de confusão entre as classes *Blight* e *Gray Leaf Spot* indica a necessidade de um conjunto de dados mais robusto e com maior distinção visual entre patologias morfológicamente semelhantes.

Conclui-se que a abordagem é promissora e representa um passo importante hacia a automação do diagnóstico fitossanitário. No entanto, para que a solução seja viável em aplicações práticas e comerciais, são necessários aprimoramentos futuros. Recomenda-se a expansão do banco de dados com imagens coletadas diretamente no campo, o ajuste fino de mais camadas da rede e a experimentação com arquiteturas alternativas para aumentar a robustez e a precisão do sistema. A continuidade desta pesquisa tem o potencial de fornecer uma ferramenta valiosa para agricultores, contribuindo para a agricultura de precisão e a segurança alimentar.

## REFERÊNCIAS

ARIZA, M.; BRUNO, O. (2024). Citado como um estudo que obteve 98,31% de acurácia na detecção de doenças em milho com a arquitetura VGG19.

CLARKE, A. C. Perfíles del Futuro (Profiles of the Future). Citado no resumo em espanhol.

EMBRAPA (2022). Publicação ou comunicado técnico da Empresa Brasileira de Pesquisa Agropecuária, utilizado para contextualizar a relevância e as perdas causadas por doenças como a ferrugem polissora e para descrever a morfologia das lesões.

FERENTINOS, K. P. (2018). Artigo científico que demonstrou alta acurácia de redes neurais na detecção de doenças em plantas, usado como referência para a eficácia do deep learning e das técnicas de aumento de dados. (Provável referência: "Deep learning models for plant disease detection and diagnosis").

HARARI, Y. N. Homo Deus: Uma Breve História do Amanhã. Citado na introdução para contextualizar o papel da inteligência artificial.

HE, K. et al. (2016). Artigo que propôs a arquitetura ResNet50, fundamental para o modelo desenvolvido no trabalho. (Referência: "Deep Residual Learning for Image Recognition").

KINGMA, D. P.; BA, J. (2014). Artigo que introduziu o otimizador Adam, utilizado no treinamento do modelo. (Referência: "Adam: A Method for Stochastic Optimization").

LECUN, Y.; BENGIO, Y.; HINTON, G. (2015). Artigo de referência que descreve os fundamentos das Redes Neurais Convolucionais (CNNs). (Provável referência: "Deep Learning", publicado na revista Nature).

MOHANTY, S. P. et al. (2016). Estudo pioneiro que demonstrou a eficácia de CNNs na classificação de doenças em plantas, servindo como base para a metodologia de divisão de dados. (Referência: "Using Deep Learning for Image-Based Plant Disease Detection").

PAN, S. J.; YANG, Q. (2010). Artigo de pesquisa sobre transfer learning (aprendizagem por transferência), técnica central aplicada no modelo ResNet50. (Provável referência: "A Survey on Transfer Learning").

PORTAL SISTEMA FAMASUL (2024). Fonte de dados sobre a produção de milho no Brasil e no Mato Grosso do Sul.

SELVARAJU, R. R. et al. (2017). Artigo que propôs a técnica Grad-CAM, utilizada para a interpretação e visualização das decisões do modelo. (Referência: "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization").

SHORTEN, C.; KHOSHGOFTAAR, T. M. (2019). Artigo de revisão sobre técnicas de aumento de dados (data augmentation), cujas recomendações foram seguidas no trabalho. (Referência: "A Survey on Image Data Augmentation for Deep Learning"). Fonte de Dados (Dataset)

KAGGLE. O conjunto de imagens utilizado para treinar, validar e testar o modelo foi o "Corn Leaf Disease Dataset", obtido do repositório Kaggle.