

Identificação de Doenças de Plantas Baseada em Visão para Sistemas Autônomos de Manejo de Culturas

Emanuelle S Gil¹, Lucas M. A. Dias¹, Alternei S. Brito¹, Felipe G. Oliveira¹

¹Instituto de Ciências Exatas e Tecnologia – Universidade Federal do Amazonas
Itacoatiara - AM, Brasil

emanuelle.gil@ufam.edu.br, lucas.dias@ufam.edu.br, alternei@ufam.edu.br,

felipeoliveira@ufam.edu.br

Abstract. *The integration of robotics in agriculture is revolutionizing traditional practices to enable intelligent plant health monitoring and disease detection. This work proposes a new approach for plant disease identification, using the ConvNeXt deep learning model to extract and classify plant visual features across species and disease types. Two datasets were used: Plant Village and Plant Pathology 2020. The proposed method achieved high accuracy, with 99.47% in Plant Village and 93.83% in Plant Pathology 2020, outperforming comparative approaches. The model was shown to be robust and scalable to independent robots, supporting more sustainable agriculture.*

Resumo. *A integração da robótica na agricultura está revolucionando práticas tradicionais para permitir o monitoramento inteligente da saúde das plantas e a detecção de doenças. Este trabalho propõe uma nova abordagem para a identificação de doenças em plantas, utilizando o modelo de aprendizado profundo ConvNeXt para extrair e classificar características visuais de plantas em diferentes espécies e tipos de doenças. Foram utilizados dois conjuntos de dados: Plant Village e Plant Pathology 2020. O método proposto obteve alta acurácia, com 99,47% no Plant Village e 93,83% no Plant Pathology 2020, superando abordagens comparativas. O modelo foi mostrado robusto e escalável para robôs independentes, apoiando uma agricultura mais sustentável.*

1. Introdução

O rápido crescimento da população global aumentou significativamente a demanda pela produção de alimentos, representando um grande desafio para a agricultura moderna. Garantir a segurança alimentar requer práticas agrícolas sustentáveis e eficientes para maximizar a produtividade das culturas e, ao mesmo tempo, minimizar o consumo de recursos. A agricultura desempenha um papel fundamental na economia global, fornecendo bens essenciais e oportunidades de emprego. No entanto, diversos desafios ameaçam a produtividade agrícola, incluindo mudanças climáticas, degradação do solo e a disseminação de doenças de plantas. Entre eles, as doenças de plantas representam uma questão crítica, pois podem levar a perdas substanciais na produtividade das culturas, afetando tanto o suprimento de alimentos quanto a estabilidade econômica.

A detecção precoce e precisa de doenças é essencial para mitigar o impacto negativo das doenças de plantas na produtividade agrícola. Tradicionalmente, a identificação

de doenças tem sido realizada manualmente por agricultores e especialistas agrícolas por meio da inspeção visual das plantas. Embora essa abordagem seja amplamente utilizada, ela consome tempo, exige muita mão de obra e está sujeita a erros humanos, especialmente em fazendas de grande porte. A necessidade de soluções automatizadas e escaláveis levou a pesquisa a abordagens tecnológicas para a detecção de doenças de plantas.

Os avanços tecnológicos permitiram o desenvolvimento de sistemas automatizados de detecção de doenças em plantas, oferecendo soluções eficientes e escaláveis para apoiar a agricultura moderna. Entre essas tecnologias, a robótica, a visão computacional e a inteligência artificial surgiram como ferramentas poderosas para automatizar o monitoramento da saúde das culturas. Ao alavancar a análise baseada em imagens e modelos de aprendizado de máquina, esses sistemas podem identificar doenças em plantas com alta precisão, reduzindo a dependência da inspeção manual e permitindo uma intervenção eficiente. A integração de plataformas robóticas com a detecção de doenças baseada em visão aumenta a escalabilidade e a eficiência dessas soluções, abrindo caminho para a agricultura de precisão.

Neste artigo, propomos uma abordagem baseada em visão para a identificação de doenças em plantas utilizando um modelo de aprendizado profundo ConvNeXt. Para isso, um modelo ConvNeXt foi empregado para aprender as características visuais nos diferentes tipos de plantas, identificando plantas saudáveis e doentes. Nos experimentos, utilizamos dois conjuntos de dados de imagens bem estabelecidos, o Plant Village e o Plant Pathology 2020. Além disso, foi realizada uma análise comparativa, comparando nossos resultados com outros algoritmos de última geração bem estabelecidos na área. Os resultados comprovam a eficácia da nossa abordagem para a identificação de doenças em plantas, destacando uma precisão significativa na operação de classificação. Também avaliamos a metodologia proposta na presença de ruído, onde os resultados demonstraram a robustez da nossa abordagem.

Nossas principais contribuições estão resumidas a seguir:

- Propomos uma abordagem baseada em Aprendizado Profundo para o monitoramento da saúde vegetal, identificando plantas doentes. O processo de classificação proposto, utilizando um modelo ConvNeXt, aprende e representa as principais características das plantas, aprimorando a detecção de doenças durante a inspeção. Essa abordagem inovadora oferece uma solução robusta para a inspeção complexa de diferentes tipos de plantas, alcançando precisão significativa mesmo na presença de ruído.
- Realizamos experimentos extensivos para a validação da nossa metodologia utilizando dois conjuntos de dados bem definidos. Esses experimentos tiveram como objetivo comparar quantitativamente modelos de aprendizado profundo com o modelo proposto, especificamente em cenários desafiadores envolvendo a presença de ruído. Os resultados fornecem insights significativos que podem orientar futuras pesquisas nessa área.

2. Objetivos

Este trabalho tem como principal objetivo desenvolver uma abordagem robusta e eficiente para a identificação automática de doenças em plantas, utilizando técnicas de apren-

dizado profundo aplicadas à visão computacional. Busca-se, por meio da arquitetura ConvNeXt, explorar a capacidade dos modelos neurais modernos em extrair e classificar características visuais relevantes, permitindo distinguir entre plantas saudáveis e doentes com elevada precisão. A proposta visa atender à crescente demanda por soluções tecnológicas que automatizem o monitoramento da saúde vegetal, minimizando a necessidade de análises manuais e ampliando a rapidez e a confiabilidade na detecção de anomalias que afetam o desenvolvimento das culturas.

3. Fundamentação Teórica

A classificação de doenças de plantas é amplamente explorada devido ao impacto das pragas na produtividade e nas perdas econômicas. Diversas abordagens têm sido propostas para mitigar esses desafios usando técnicas de aprendizado profundo e visão computacional [Pendhari et al. 2023] [Hashemifar and Zakeri-Nasrabadi 2024].

Modelos baseados em Redes Neurais Convolucionais (CNNs) são amplamente aplicados nesta área, destacando a variabilidade de modelos não pré-treinados para cenários complexos, embora os desafios em ambientes não controlados sejam ressaltados. O trabalho de [Belmir et al. 2023] propõe um modelo de CNN para classificação de doenças em folhas de plantas, utilizando o conjunto de dados PlantVillage, que abrange 38 classes de doenças em 14 culturas. O objetivo é automatizar o diagnóstico precoce de doenças, substituindo métodos manuais e subjetivos, com foco em aumentar a acurácia e a eficiência na agricultura. A arquitetura da CNN é composta por três blocos convolucionais com 32, 64 e 128 filtros, além de camadas de pooling e dropout, alcançando 94,33% de acurácia, superando modelos como YOLOv8 e DenseNet-77.

Em [Lakshmanarao et al. 2021] é proposto um modelo baseado em CNN para detecção e classificação de doenças em plantas, utilizando o conjunto de dados público PlantVillage, com foco em folhas de batata, pimentão e tomate. As imagens foram redimensionadas para 64x64 pixels, alcançando acurácias de 98,3%, 98,5% e 95%, respectivamente. O trabalho destaca a eficácia de modelos especializados, embora os desafios em cenários complexos sejam enfatizados. A baixa resolução das imagens pode comprometer detalhes sutis. O trabalho de [Kolakaluru et al. 2023] propõe um sistema baseado em CNN para identificação automatizada de doenças e pragas em culturas, com o objetivo de aprimorar o Manejo Integrado de Pragas (MIP) e reduzir perdas econômicas no setor agrícola. Os resultados apresentaram 97,5% de acurácia na identificação de pragas, integrando técnicas de imagem hiperespectral para superar desafios como camuflagem, superando técnicas como o YOLOv5.

Além dos modelos tradicionais de CNN, o aprendizado por transferência tem se mostrado uma estratégia eficaz para lidar com os desafios da classificação de doenças em plantas. O trabalho de [Pendhari et al. 2023] compara o desempenho de modelos de aprendizado por transferência, como MobileNetV2, VGG16, VGG19, NASNet e Inception-ResNet, na detecção de doenças em plantas, utilizando o conjunto de dados New Plant Diseases, com 31 classes de doenças em 9 culturas. O objetivo é identificar o modelo mais eficaz para aplicações práticas na agricultura. Os modelos pré-treinados foram ajustados com modificações nas camadas densas e funções de ativação. O artigo oferece uma análise valiosa dos modelos de transferência para detecção de doenças, destacando o VGG19 como o mais preciso, com 94% de acurácia, embora menos eficiente.

Sua principal contribuição é a comparação estruturada entre arquiteturas, útil para pesquisadores e desenvolvedores que precisam equilibrar precisão e recursos computacionais.

Complementando, [Pavan et al. 2023] propôs um modelo baseado em aprendizado por transferência com a arquitetura EfficientNetb5 para classificação de doenças em folhas de plantas, utilizando um conjunto de dados do Kaggle com 38 classes e 80.000 imagens. O modelo foi comparado com as arquiteturas Inception-v3 e VGG16, alcançando 99,20% de acurácia, superando os demais e demonstrando equilíbrio entre precisão e generalização. No entanto, a falta de detalhes sobre o balanceamento do conjunto de dados e validação externa limita a interpretação da robustez do modelo.

Modelos híbridos surgiram como alternativas promissoras ao combinar as vantagens de diferentes arquiteturas para melhorar a acurácia na classificação de doenças em plantas. O trabalho de [Hashemifar and Zakeri-Nasrabadi 2024] propõe um modelo híbrido que combina o Vision Transformer (ViT) e o VGG-16 para detecção de doenças em plantas, utilizando o conjunto de dados PlantVillage. O objetivo é superar as limitações de modelos baseados exclusivamente em CNNs ou ViTs, equilibrando a captação de características locais e globais. Para isso, a abordagem combina a extração hierárquica de características do VGG-16 com a capacidade de captura de contexto global do Pyramid ViT (PVT). O modelo alcançou 98,51% de acurácia, superando outros modelos combinados. A junção do VGG-16 com o PVT aborda questões como a necessidade de grandes conjuntos de dados e a ausência de contexto global. No entanto, o custo computacional e o desbalanceamento entre classes ainda são desafios a serem enfrentados.

Outros estudos exploraram métodos comparativos para avaliar o desempenho de diferentes modelos. O trabalho de [Melese and Yayeh 2023] propõe modelos híbridos que combinam CNNs com métodos clássicos de aprendizado de máquina, como Máquinas de Vetores de Suporte (SVM), Random Forest e Árvore de Decisão, para identificação e classificação de doenças em folhas de soja. A ideia é aproveitar a capacidade de extração de características profundas das CNNs e a robustez dos classificadores tradicionais, visando melhorar a acurácia e a eficiência no diagnóstico de doenças agrícolas. Três arquiteturas híbridas foram desenvolvidas: CNN-SVM, CNN-RF e CNN-DT, substituindo a camada softmax da CNN por classificadores tradicionais. A CNN-RF alcançou 78,41% de acurácia na validação. Este trabalho abre caminho para futuras pesquisas que integrem técnicas de deep learning com algoritmos clássicos em cenários mais diversificados.

O trabalho de [Hanif et al. 2024] conduziu um estudo comparando ResNet, Inception-v3 e SVM utilizando o CIFAR-10, um conjunto de dados genérico amplamente usado para classificação de imagens. Os resultados mostraram que a ResNet obteve 99,81% de precisão, superando o Inception-v3, que obteve 87%, enquanto o SVM apresentou o melhor desempenho, com 99,88%. [Ermolaeva 2024] utilizou aprendizado de máquina com 70.000 imagens de 38 classes, treinando modelos como ResNet-9, ResNet-18 e VGG-16. O ResNet-9 se destacou, alcançando 99,23% de precisão em apenas cinco épocas.

Além disso, abordagens que combinam deep learning com aprendizado de máquina têm se mostrado eficazes em cenários específicos. [Jain et al. 2023] comparou Regressão Logística, Redes Neurais Convolucionais e métodos híbridos de deep learning para classificar doenças em folhas de milho, utilizando os conjuntos de dados PlantVil-

lage e PlantDoc. O método proposto alcançou 93,2% de acurácia, destacando o potencial das técnicas híbridas, apesar dos desafios enfrentados em cenários do mundo real.

Por fim, [Bhargava et al. 2024] revisou o uso de inteligência artificial e visão computacional na detecção de doenças em plantas, destacando o papel de modelos como ResNet e EfficientNet, bem como o uso de dados RGB e hiperespectrais para alcançar alta precisão. Os autores também enfatizaram os desafios relacionados à necessidade de conjuntos de dados e à generalização para condições reais de campo, sugerindo que avanços futuros devem focar em técnicas robustas e aplicáveis no ambiente agrícola.

Esta visão geral destaca os avanços na área, mas também revela lacunas significativas em termos de generalização para condições reais de cultivo, uma questão que este estudo aborda de forma inovadora. Embora os trabalhos revisados apresentem abordagens eficazes para a identificação de doenças em plantas, muitos enfrentam dificuldades na aplicação de seus resultados em cenários do mundo real. Este estudo busca superar essas limitações ao propor uma abordagem robusta e eficiente, que integra técnicas modernas de deep learning e visão computacional para sistemas autônomos de manejo agrícola.

4. Metodologia

Este artigo aborda o problema do gerenciamento da saúde das plantas utilizando a arquitetura de deep learning ConvNeXt. A metodologia proposta explora a identificação de doenças em plantas por meio da extração e classificação de características visuais, com ênfase na eficiência e precisão do modelo. Os detalhes da arquitetura ConvNeXt, assim como a configuração dos experimentos realizados, serão apresentados nas subseções seguintes.

4.1. Arquitetura ConvNeXt para Identificação de Doenças em Plantas

ConvNeXt é uma arquitetura baseada em redes neurais convolucionais (CNNs), projetada para combinar os avanços das arquiteturas tradicionais de CNN com ideias inspiradas em Transformers. Desenvolvida com o objetivo de melhorar o desempenho e a eficiência computacional, ConvNeXt introduz otimizações como convoluções profundas, camadas de normalização adaptativas e uma organização hierárquica das camadas para aprimorar a capacidade de extração de características. Entre os principais diferenciais do ConvNeXt estão o uso de convoluções separáveis e a aplicação de camadas de normalização que eliminam a necessidade de camadas como a Batch Normalization, reduzindo a complexidade e aumentando a eficiência. Essas melhorias tornam o ConvNeXt uma escolha competitiva para tarefas complexas de classificação de imagens, como a detecção de doenças em plantas [Liu et al. 2022]. Mais detalhes sobre a arquitetura ConvNeXt podem ser encontrados na Figura 1.

A arquitetura do ConvNeXt consiste em:

4.1.1. Bloco Residual do ConvNeXt

Cada bloco residual no ConvNeXt segue uma estrutura de gargalo invertido, semelhante à do Transformer [Liu et al. 2022]. A saída do bloco é dada por:

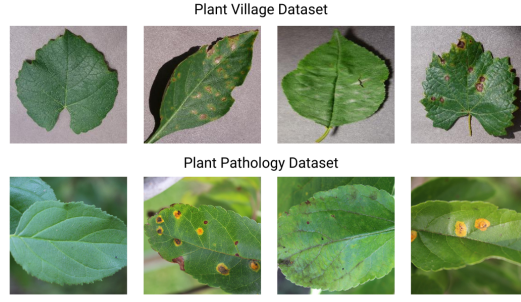


Figura 1. Proposed ConvNeXt architecture.

$$y = x + \alpha \cdot Block(x), \quad (1)$$

onde x é a entrada do bloco, $Block(x)$ representa as operações realizadas no bloco (convoluções, normalização, ativação) e α é um parâmetro escalar.

4.1.2. Gargalo Invertido

Inspirado pelo MobileNetV2, o bloco expande a dimensionalidade dos canais antes de aplicar a convolução depthwise e, em seguida, a reduz novamente [Sandler et al. 2018]. A estrutura é:

$$z = Proj_{\uparrow}(x), \quad (2)$$

$$z' = DWConv_{k \times k}(z), \quad (3)$$

$$y = Proj_{\downarrow}(z') + x, \quad (4)$$

onde $Proj_{\uparrow}$ é a projeção linear (1×1 conv) que aumenta os canais, $DWConv_{k \times k}$ é a convolução depthwise com kernel $k \times k$, e $Proj_{\downarrow}$ é a projeção linear que reduz os canais de volta à sua dimensão original.

4.1.3. Convoluções com Kernel Grande

O ConvNeXt utiliza kernels grandes (7×7) em convoluções depthwise, contrariando a tendência de kernels pequenos (3×3) em CNNs tradicionais [Liu et al. 2022]. Isso aumenta o campo receptivo para capturar contextos globais, de forma semelhante aos Transformers. A operação é:

$$z' = W_{k \times k} * z, \quad (5)$$

onde $W_{k \times k}$ é o kernel de convolução $k \times k$ ($k = 7$).

4.1.4. LayerNorm

Substitui a BatchNorm, comum em CNNs, pela LayerNorm, utilizada em Transformers [Lei Ba et al. 2016]. Para um tensor de entrada $x \in \mathbb{R}^{B \times C \times H \times W}$, a LayerNorm é aplicada aos canais (C):

$$\hat{x} = \frac{x - \mu}{\sigma} \cdot \gamma + \beta, \quad (6)$$

onde μ e σ são a média e o desvio padrão calculados sobre C, H, W , e γ e β são parâmetros aprendidos (escala e viés).

4.1.5. Ativação GELU

Substitui a ReLU pela função GELU [Hendrycks and Gimpel 2016], utilizada em Transformers, definida como:

$$GELU(x) = x \cdot \Phi(x), \quad (7)$$

onde $\Phi(x)$ é a função de distribuição acumulada (CDF) da normal padrão. Aproximadamente:

$$GELU(x) \approx 0.5x \left(1 + \tanh \left(\sqrt{\frac{2}{\pi}} (x + 0.044715x^3) \right) \right). \quad (8)$$

4.1.6. LayerScale

Técnica introduzida no ConvNeXt para estabilizar o treinamento [Touvron et al. 2021]. Multiplica a saída do bloco residual por um vetor diagonal $\alpha \in \mathbb{R}^C$:

$$y = x + \alpha \odot Block(x), \quad (9)$$

onde \odot denota a multiplicação elemento a elemento. Inicialmente, α é definido com um valor pequeno.

4.1.7. Stochastic Depth

Uma técnica que desativa aleatoriamente blocos durante o treinamento, prevenindo o overfitting [Huang et al. 2016]. Para um bloco na camada l , a saída é:

$$y = x + b_l \cdot Block(x), \quad (10)$$

onde $b_l \sim Bernoulli(p_l)$ é uma variável aleatória binária e p_l é a probabilidade de sobrevivência do bloco.

O ConvNeXt é uma arquitetura de deep learning fundamentada em sua capacidade única de combinar a eficiência comprovada das redes convolucionais clássicas (CNNs) com avanços inspirados nos Vision Transformers (ViTs), resultando em um modelo robusto, escalável e adaptável para diversas tarefas de visão computacional. Sua superioridade decorre da preservação de vieses indutivos das convoluções, os quais mantêm a robustez a variações espaciais, algo crucial para tarefas de detecção e segmentação.

Ao adaptar estratégias fundamentais dos ViTs ao paradigma convolucional, a arquitetura ConvNeXt se destaca em benchmarks como ImageNet-C e ImageNet-A, por meio do uso de LayerNorm e kernels grandes. Isso demonstra que as convoluções continuam sendo essenciais na era dos Transformers, oferecendo um modelo preciso, eficiente e prático para várias aplicações.

5. Resultados e Discussão

5.1. Configurações experimentais

5.1.1. Conjuntos de Dados

Para avaliar a abordagem proposta, adotamos dois conjuntos de dados desafiadores. O Plant Village [G. and J. 2019] contém 55.448 imagens de 13 tipos de folhas de plantas saudáveis e doentes, distribuídas em 39 classes. O conjunto de dados Plant Pathology 2020 [Kaeser-Chen et al. 2020] é composto por 4 classes e inclui 3.651 imagens de folhas de macieiras afetadas por doenças como Apple Scab, Ferrugem e Cedar Apple Rust, além de folhas saudáveis.

Para aumentar a robustez e a capacidade de generalização do modelo de aprendizado proposto, o conjunto de dados foi artificialmente expandido com diversas variações por meio de técnicas de aumento de dados (data augmentation). Para isso, foram utilizados parâmetros variados de recorte, mistura e variação de cores (color jittering), gerando imagens mais desafiadoras. Além disso, as imagens foram redimensionadas para 224×224 pixels para se adequar às dimensões esperadas pelo tensor de entrada dos modelos. A Figura 2 apresenta amostras de folhas de plantas dos conjuntos de dados utilizados.

5.1.2. Detalhes de Implementação

Utilizamos o framework PyTorch em um computador Dell com processador Intel® Xeon™ Silver 4114 2.20GHz, 128 GB de memória principal DDR4-2133 e uma GPU NVIDIA® GeForce® RTX A4000 com 16 GB de memória GDDR6 para a nossa abordagem. A fase de treinamento do modelo ConvNeXt proposto envolveu o uso de Grid Search para otimização de hiperparâmetros (taxa de aprendizado, tamanho do batch, número de épocas e taxa de dropout), visando alta acurácia. As abordagens baseadas em aprendizado utilizaram um conjunto de imagens de entrada, divididas em conjuntos de treinamento e teste, garantindo a ausência de sobreposição por meio do protocolo de validação cruzada com 5 folds.

5.2. Avaliação da Identificação de Doenças em Plantas

Este experimento avalia a acurácia do método proposto para identificação de doenças em plantas. Oito estratégias de classificação diferentes são avaliadas: *i*) abordagem

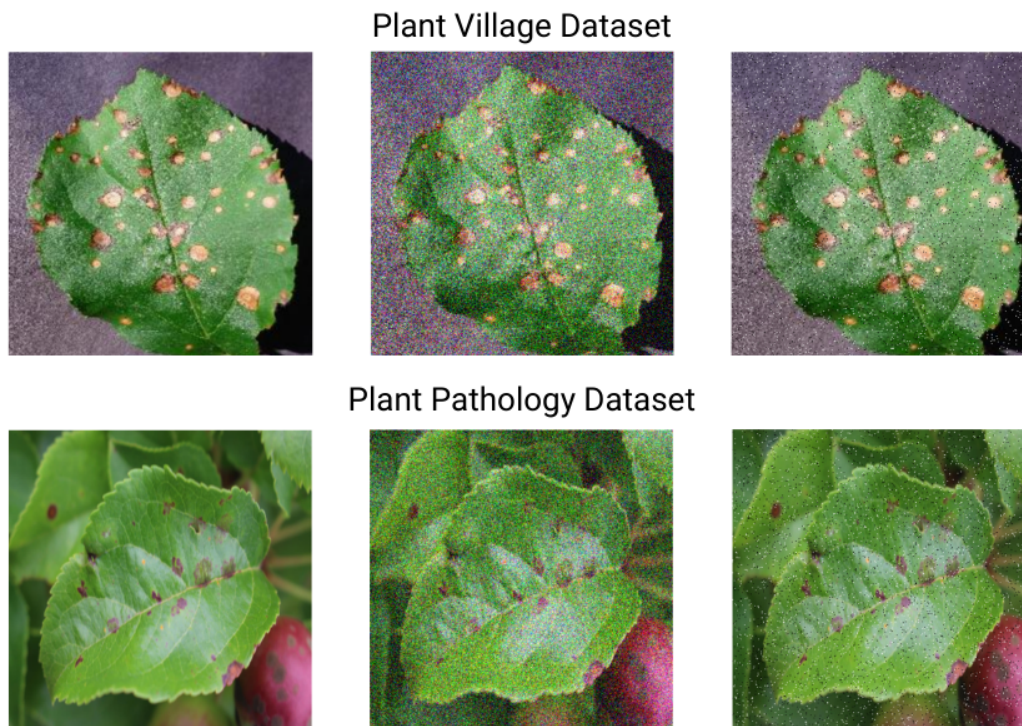


Figura 2. Exemplos de folhas de plantas que compõem os conjuntos de dados, destacando folhas saudáveis e doentes.

ConvNeXt (proposta); *ii*) Swin Transformer V2; *iii*) MobileViT; *iv*) ConvFormer; *v*) EfficientFormer; *vi*) BeiT; *vii*) DeiT; e *viii*) MLP Mixer. Técnicas de avaliação baseadas em deep learning são empregadas devido ao seu alto desempenho em problemas de classificação e segmentação, ainda representando o estado da arte na área, especialmente em aplicações voltadas à agricultura de precisão [Fahim-Ul-Islam et al. 2024] [Mu et al. 2024] [Yu et al. 2024] [Ji et al. 2024] [Ahmed et al. 2024] [Ma et al. 2023] [Li and Tanone 2022].

Os resultados indicam que o modelo ConvNeXt proposto supera as demais técnicas comparadas de deep learning, conforme mostrado na Tabela I, considerando tanto a acurácia quanto o desvio padrão. O ConvNeXt combina eficiência computacional, alta precisão e simplicidade arquitetural. Ele utiliza convoluções depthwise para substituir o mecanismo de self-attention dos Transformers, reduzindo a carga computacional. O Gargalo Invertido (Inverted Bottleneck) expande os canais antes da convolução espacial, otimizando os recursos. A integração de vieses indutivos e técnicas modernas, como kernels grandes e a ativação GELU, permite capturar o contexto global de forma mais eficiente e com custo linear. Além disso, o ConvNeXt estabiliza o treinamento e melhora a convergência em redes profundas. Sem a necessidade de módulos especializados e com uma arquitetura totalmente convolucional, ele simplifica a implementação e é eficiente para aprendizado por transferência, mantendo alta acurácia sem necessidade de extensos ajustes finos [Liu et al. 2022].

Tabela 1. Resultados da Identificação de Doenças em Plantas como um problema de classificação. Este experimento apresenta a acurácia dos métodos de deep learning ConvNeXt (proposto), Swin Transformer V2, MobileViT, ConvFormer, EfficientFormer, BeiT, DeiT e MLP-Mixer.

Método	Acurácia no Plant Village	Acurácia no Plant Pathology
Swin Transformer V2 [20]	99.06 \pm 0.05	90.58 \pm 5.25
MobileViT [21]	98.40 \pm 0.08	91.75 \pm 0.75
ConvFormer [22]	99.28 \pm 0.08	92.42 \pm 0.49
EfficientFormer [23]	99.15 \pm 0.09	90.06 \pm 1.11
BeiT [24]	98.89 \pm 0.09	90.03 \pm 0.35
DeiT [25]	99.04 \pm 0.10	92.21 \pm 0.41
MLP-Mixer [26]	98.82 \pm 0.16	89.44 \pm 0.71
ConvNeXt (Proposto)	99.47 \pm 0.06	93.83 \pm 0.67

5.3. Avaliação da Identificação de Doenças em Plantas na Presença de Ruído

Neste experimento, avaliamos a robustez da abordagem proposta para identificação de doenças em plantas sob condições com ruído. Para isso, adicionamos ruído Gaussiano e ruído Sal e Pimenta às imagens de folhas do conjunto de dados Plant Pathology e comparamos o desempenho de diversas técnicas de deep learning, incluindo nossa solução baseada no modelo ConvNeXt. Exemplos de imagens de plantas doentes com ruído são ilustrados na Figura 3. A Figura 3a apresenta uma imagem de planta doente sem ruído. A Figura 3b mostra uma imagem com ruído Gaussiano, considerando densidade de ruído de 0.02. A Figura 3c apresenta uma imagem com ruído Sal e Pimenta, também com densidade de ruído de 0.02.

Esse ruído simula desafios reais na aquisição de imagens para identificação de doenças em plantas. Durante a fase de treinamento, foram utilizadas imagens sem ruído, enquanto a fase de teste envolveu imagens com ruído adicionado, utilizando um processo de aprendizado por transferência.

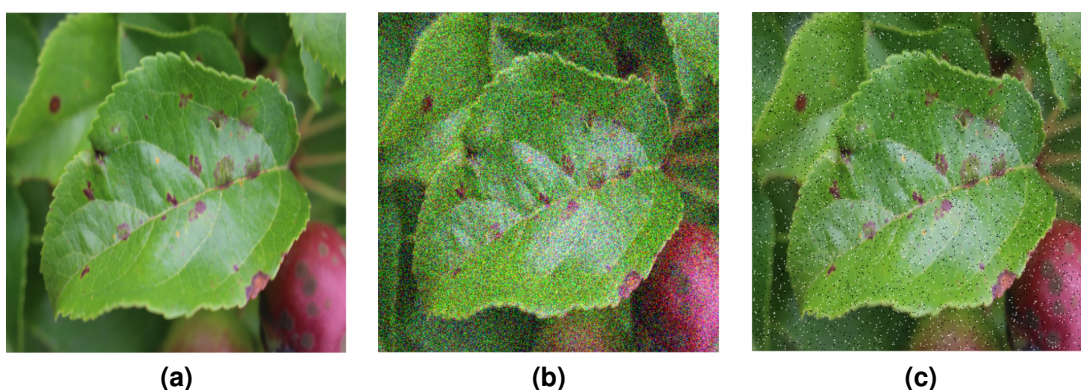


Figura 3. Exemplos de imagens de plantas com e sem ruído: Figura 3a sem ruído; Figura 3b com ruído Gaussiano; Figura 3c com ruído Sal e Pimenta. Ambas as imagens com ruído consideram densidade de ruído igual a 0,02.

Os resultados deste experimento indicam que o modelo ConvNeXt proposto apresenta melhor desempenho mesmo sob condições com ruído, conforme mostrado nas tabelas a seguir. A Tabela 2 apresenta os resultados de classificação para o ruído Gaussiano,

incluindo acurácia e desvio padrão, utilizando o conjunto de dados Plant Pathology. A Tabela 3 apresenta os resultados para o ruído Sal e Pimenta, também com acurácia e desvio padrão, para o mesmo conjunto de dados. Esses resultados demonstram que a abordagem proposta supera os demais métodos de classificação, evidenciando sua robustez.

Tabela 2. Resultados da avaliação de robustez para o problema de identificação de doenças em plantas, considerando a presença de ruído Gaussiano.

Ruído Gaussiano			
Densidade de ruído	0,005	0,01	0,02
ConvFormer [Yu et al. 2024]	91,32 ± 0,33	87,91 ± 1,30	81,10 ± 1,32
DeiT [Ma et al. 2023]	92,31 ± 0,50	90,39 ± 0,65	84,27 ± 1,52
MobileViT [Mu et al. 2024]	89,23 ± 0,96	78,98 ± 0,57	44,61 ± 3,12
ConvNeXt (Proposto)	93,72 ± 0,38	92,35 ± 0,43	88,26 ± 1,17

Tabela 3. Resultados da avaliação de robustez para o problema de identificação de doenças em plantas, considerando a presença de ruído Sal e Pimenta.

Ruído Sal e Pimenta			
Densidade de ruído	0,005	0,01	0,02
ConvFormer [22]	81,86 ± 2,93	71,37 ± 4,68	63,08 ± 3,43
DeiT [25]	89,46 ± 0,98	85,63 ± 0,31	75,27 ± 1,06
Mobile ViT [21]	67,35 ± 1,57	62,29 ± 4,53	54,42 ± 2,40
ConvNeXt (Proposto)	90,94 ± 0,62	86,45 ± 0,85	76,99 ± 4,40

Para esses experimentos, comparamos nossa abordagem com os modelos DeiT, ConvFormer e MobileViT, considerando seu alto desempenho no experimento anterior. Utilizamos o conjunto de dados Plant Pathology, por conter imagens mais representativas de um ambiente real de doenças em plantas.

Os experimentos demonstraram que a arquitetura ConvNeXt superou todas as abordagens comparadas, tanto em condições normais quanto na presença de ruídos artificiais, como ruído Gaussiano e ruído Sal e Pimenta. Esses resultados indicam não apenas a superioridade do ConvNeXt em termos de acurácia, mas também sua resiliência a variações e distúrbios nos dados de entrada. Essa robustez é essencial para a aplicação prática do modelo em ambientes agrícolas reais, nos quais as imagens capturadas estão sujeitas a diferentes condições de iluminação, variações na qualidade da câmera e interferências externas. Diante desses achados, o ConvNeXt destaca-se como uma solução promissora para sistemas autônomos de monitoramento da saúde de plantas. A seguir, discutiremos em detalhes as implicações desses resultados e possíveis direções futuras para o aprimoramento do modelo.

6. Conclusão

Neste estudo, demonstramos que a arquitetura ConvNeXt é altamente eficaz para a identificação de doenças em plantas, superando abordagens comparativas mesmo sob condições adversas com a introdução de ruído artificial. A alta acurácia alcançada nos conjuntos de dados Plant Village (99,47%) e Plant Pathology 2020 (93,83%) reforça a capacidade do modelo de generalizar em diferentes cenários e espécies vegetais.

Além disso, a resiliência do modelo a perturbações visuais destaca sua adequação para aplicações em campo, onde variações de iluminação e qualidade de imagem representam desafios reais. A integração de sistemas autônomos para o monitoramento da saúde das plantas é um passo fundamental para o avanço da agricultura de precisão, possibilitando maior produtividade e redução no desperdício de recursos. Os resultados obtidos neste trabalho evidenciam o potencial do ConvNeXt como uma solução escalável e eficiente para esses sistemas, contribuindo para o progresso da robótica agrícola e para a implementação de práticas agrícolas mais sustentáveis.

Como trabalho futuro, planejamos submeter um artigo para periódico apresentando os resultados complementares obtidos em cenários mais desafiadores. Além disso, investigações futuras incluirão a avaliação de novos modelos de deep learning e a adaptação da metodologia para identificação de doenças em diferentes tipos de culturas. Por fim, pretendemos aprimorar a abordagem de identificação de doenças em cenários reais, considerando diferentes condições de iluminação e o uso de distintos veículos robóticos.

Referências

- Ahmed, S. T., Barua, S., Fahim-Ul-Islam, M., and Chakrabarty, A. (2024). Enhancing precision in rice leaf disease detection: A transformer model approach with attention mapping. In *2024 Int. Conf. on Adv. in Comp., Com., Elect., and Smart Sys. (iCAC-CESS)*, pages 1–6.
- Belmir, M., Difallah, W., and Ghazli, A. (2023). Plant leaf disease prediction and classification using deep learning. In *2023 Int. Conference on Decision Aid Sciences and Applications (DASA)*, pages 536–540. IEEE.
- Bhargava, A., Shukla, A., Goswami, O., Alsharif, M. H., Uthansakul, P., and Uthansakul, M. (2024). Plant leaf disease detection, classification and diagnosis using computer vision and artificial intelligence: A review. *IEEE Access*.
- Ermolaeva, A. D. (2024). Plant disease detection using small convolutional neural networks. In *2024 Conference of Young Researchers in Electrical and Electronic Engineering (ElCon)*, pages 10–12. IEEE.
- Fahim-Ul-Islam, M., Chakrabarty, A., Ahmed, S. T., Rahman, R., Kwon, H. H., and Jallil Piran, M. (2024). A comprehensive approach toward wheat leaf disease identification leveraging transformer models and federated learning. *IEEE Access*, 12:109128–109156.
- G., G. and J., A. P. (2019). Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Computers Electrical Engineering*, 76:323–338.
- Hanif, M. A., Zim, M. K. I., and Kaur, H. (2024). Resnet vs inception-v3 vs svm: A comparative study of deep learning models for image classification of plant disease detection. In *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, volume 2, pages 1–6. IEEE.
- Hashemifar, S. and Zakeri-Nasrabadi, M. (2024). Deep identification of plant diseases. In *2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP)*, pages 1–6. IEEE.
- Hendrycks, D. and Gimpel, K. (2016). Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*.
- Huang, G., Sun, Y., Liu, Z., Sedra, D., and Weinberger, K. Q. (2016). Deep networks with stochastic depth. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14*, pages 646–661. Springer.
- Jain, S., Jaidka, P., and Jain, V. (2023). Plant leaf disease classification using deep learning based hybrid approach. In *2023 Int. Conf. on Com., Security and Artificial Intelligence (ICCSAI)*, pages 383–387. IEEE.
- Ji, W., Zhang, T., Xu, B., and He, G. (2024). Apple recognition and picking sequence planning for harvesting robot in a complex environment. *Journal of Agricultural Engineering*, 55(1).

- Kaeser-Chen, C., Pathology, F., Maggie, and Dane, S. (2020). Plant pathology 2020 - fgvc7. <https://kaggle.com/competitions/plant-pathology-2020-fgvc7>. Kaggle.
- Kolakaluru, H., Vishal, T., Chandu, M. P., Harshini, M., Vignesh, T., and Padyala, V. V. P. (2023). Crop disease identification using convolutional neural network. In *2023 International Conference on Inventive Computation Technologies (ICICT)*, pages 366–369. IEEE.
- Lakshmanarao, A., Babu, M. R., and Kiran, T. S. R. (2021). Plant disease prediction and classification using deep learning convnets. In *2021 Int. Conf. on Art. Intel. and Machine Vision (AIMV)*, pages 1–6. IEEE.
- Lei Ba, J., Kiros, J. R., and Hinton, G. E. (2016). Layer normalization. *ArXiv e-prints*, pages arXiv-1607.
- Li, L.-H. and Tanone, R. (2022). Mlp-mixer approach for corn leaf diseases classification. In Nguyen, N. T., Tran, T. K., Tukayev, U., Hong, T.-P., Trawiński, B., and Szczerbicki, E., editors, *Intelligent Information and Database Systems*, pages 204–215, Cham. Springer Nature Switzerland.
- Liu, Z., Mao, H., , C.-Y., Feichtenhofer, C., Darrell, T., and Xie, S. (2022). A convnet for the 2020s. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11976–11986.
- Ma, C., Hu, Y., Liu, H., Huang, P., Zhu, Y., and Dai, D. (2023). Generating image descriptions of rice diseases and pests based on deit feature encoder. *Applied Sciences*, 13(18).
- Melese, T. and Yayeh, Y. (2023). Hybrid deep-machine learning based performance comparison for soybean plant disease identification. In *2023 International Conference on Information and Communication Technology for Development for Africa (ICT4DA)*, pages 96–101.
- Mu, H., Zeng, X., Liu, W., Sun, J., and Zhang, Y. (2024). Improved mobile-vit model and its application in rolling bearing fault diagnosis. In *2024 6th International Conference on System Reliability and Safety Engineering (SRSE)*, pages 308–315.
- Pavan, C. H. T., Sadha, C. K., Harshini, P., Annepu, V., Bagadi, K., and Chirra, V. R. R. (2023). Plant leaf disease classification using transfer learning using efficientnetb5. In *2023 International Conference on Next Generation Electronics (NEleX)*, pages 1–6. IEEE.
- Pendhari, H., Virkar, R., Edakkalathur, B., and Jadhav, A. (2023). A comparative study on algorithms for plant disease detection using transfer learning. In *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*, pages 528–533. IEEE.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE CVPR*, pages 4510–4520.

- Touvron, H., Cord, M., Sablayrolles, A., Synnaeve, G., and Jégou, H. (2021). Going deeper with image transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 32–42.
- Yu, W., Si, C., Zhou, P., Luo, M., Zhou, Y., Feng, J., Yan, S., and Wang, X. (2024). Metaformer baselines for vision. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 46(2):896–912.