REDES NEURAIS CONVOLUCIONAIS PROFUNDAS NA DETECÇÃO DE ERVAS DANINHAS

Thiago Silva Sousa[1]
1Mestrando. Rod. Washington Luís, s/n - Monjolinho, São Carlos - SP, 13565-905. UFSCar - Universidade Federal de São Carlos, thiago.sousa@ufscar.com.br

RESUMO

Este estudo enfatiza a importância da detecção de ervas daninhas na cultura de cana-de-açúcar para aumentar a produtividade e promover a sustentabilidade, especialmente no Brasil. Após avaliar métodos de Visão Computacional e Inteligência Artificial, a arquitetura de rede neural SegFormer mostrou-se superior, alcançando uma acurácia de 95%, coeficiente de Dice de 0,95 e baixa função de perda, tornando-a a melhor escolha para detecção de ervas daninhas. A SegFormer se destacou por sua precisão e eficiência em processar imagens de drones, além de apresentar um tempo de treinamento otimizado. A pesquisa conclui que essa solução oferece uma abordagem robusta para monitorar e controlar ervas daninhas, contribuindo para a redução do uso de herbicidas e o aumento da eficiência agrícola.

Palavras-chave **—** Detecção de Ervas Daninhas, Redes Neurais Convolucionais, Agricultura de Precisão, Imagem Aérea

ABSTRACT

This study highlights the importance of weed detection in sugarcane cultivation to boost productivity and promote sustainability, particularly in Brazil. After assessing methods in Computer Vision and Artificial Intelligence, the SegFormer neural network architecture proved superior, achieving 95% accuracy, a Dice coefficient of 0.95, and low loss function, making it the best choice for weed detection. SegFormer stood out for its precision and efficiency in processing drone images, as well as its optimized training time. The study concludes that this solution offers a robust approach for weed monitoring and control, supporting reduced herbicide use and increased agricultural efficiency.

Key words **—** Weed Detection, Convolutional Neural Networks, Precision Agriculture, Aerial Imagery..

1. INTRODUÇÃO

A cana-de-açúcar é uma das culturas agrícolas mais importantes do mundo, desempenhando um papel crucial na produção de açúcar, etanol e diversos outros subprodutos. Globalmente, é cultivada em mais de 120 países, sendo uma das principais fontes de açúcar e energia renovável, além de contribuir significativamente para as economias de muitas nações tropicais e subtropicais [1].

É válido destacar que o Brasil é o maior produtor mundial de cana-de-açúcar, responsável por cerca de 40% da produção global [8], em que a cultura desempenha um papel fundamental na economia brasileira, gerando milhões de empregos diretos e indiretos, além de ser uma das principais fontes de receita de exportação do país [6]. Além disso, o Brasil é líder na produção de etanol a partir da cana-de-açúcar, utilizando-o como um biocombustível sustentável que reduz a dependência de combustíveis fósseis e contribui para a redução das emissões de Gases de Efeito Estufa (GEE) [7].

A importância dessa cultura no Brasil também se reflete na sua influência sobre o desenvolvimento de tecnologias agrícolas avançadas, voltadas para o aumento da produtividade e sustentabilidade. Logo, a detecção eficiente de ervas daninhas, por exemplo, é crucial para manter altos níveis de produtividade, e o uso de Redes Neurais Convolucionais (CNNs) para esse fim representa um avanço significativo na modernização da agricultura brasileira [4].

Portanto, a detecção de ervas daninhas nessa cultura é uma prática atualmente essencial, haja vista que essas competem com a cana-de-açúcar por recursos vitais, como água, luz e nutrientes, resultando em perdas significativas de produtividade quando não são controladas de maneira eficaz [2]. Tradicionalmente, o manejo dessas plantas invasoras é realizado através de métodos manuais ou da aplicação de herbicidas, ambos com limitações consideráveis em termos de custo, tempo e impacto ambiental [4].

Com o avanço das tecnologias de Inteligência Artificial (IA), as CNNs emergem como uma solução promissora para a detecção automatizada de ervas daninhas, devido à sua habilidade em reconhecer padrões complexos em grandes volumes de dados. Essas redes têm se destacado na Visão Computacional por sua capacidade de identificar e classificar imagens com alta precisão [5], possibilitando uma detecção mais rápida e precisa em extensas áreas de cultivo [3]. Neste contexto, o presente estudo tem como objetivo avaliar a eficácia da aplicação de CNNs para a identificação de ervas daninhas na cultura de cana-de-açúcar.

Neste contexto, o presente estudo avalia a eficácia da aplicação de CNNs para a identificação de ervas daninhas na cultura da cana-de-açúcar, com o objetio de selecionar a arquitetura de CNN mais eficaz para automatizar a detecção de plantas daninhas na cultura da cana-de-açúcar, minimizando a necessidade de intervenção humana.

2. MATERIAL E MÉTODOS

Inicialmente, foi realizado um estudo para investigar como a literatura tem abordado o uso de Redes Neurais Convolucionais (CNNs) em imagens de drones na agricultura de precisão. O objetivo desse levantamento bibliográfico foi identificar as abordagens, técnicas e aplicações mais comuns na área, além de compreender as lacunas de pesquisa e os desafios ainda presentes. Para conduzir o levantamento, foram selecionadas bases de dados acadêmicas reconhecidas, como IEEE Xplore, ScienceDirect, SpringerLink e Scopus. A pesquisa abrangeu artigos publicados nos últimos 10 anos, garantindo que as fontes fossem atuais e relevantes. Os termos de busca utilizados incluíram combinações como "CNNs", "drone imagery", "precision agriculture", e "weed detection". Foram analisados 16 artigos sobre o tema da detecção de ervas daninhas utilizando CNNs em imagens de drones.

As imagens utilizadas no estudo são multiespectrais, com um tamanho de pixel de 3 cm, e foram ortorretificadas para corrigir distorções geométricas causadas pela topografia do terreno e pela geometria da câmera do drone. A ortorretificação foi realizada utilizando um Modelo Digital de Elevação (MDE), que ajustou cada pixel da imagem com base nas variações topográficas. Esse processo corrigiu distorções de perspectiva e parallax, resultando em uma imagem alinhada com as coordenadas geográficas reais, o que garantiu a precisão geométrica necessária para as análises [5].

Além disso, foi realizado um tratamento radiométrico nas imagens com o objetivo de corrigir as variações na intensidade da luz, garantindo a consistência das informações. Esse tratamento envolveu a calibração do sensor para corrigir variações na resposta espectral do equipamento, a correção atmosférica para remover os efeitos da atmosfera na radiação refletida pela superfície do terreno, e ajustes de brilho e contraste para melhorar a qualidade da imagem. Esse procedimento foi essencial para garantir que as imagens fossem adequadas para o treinamento e a avaliação dos modelos de aprendizado de máquina, como as redes neurais convolucionais [7].

O método de anotação das imagens foi realizado de forma manual, utilizando uma técnica desenvolvida pelo autor. O processo de anotação envolveu a classificação das áreas das imagens como "daninhas" ou "não daninhas", levando em consideração as características visuais das plantas. Para as daninhas, as anotações foram feitas com base em suas características geométricas, como o formato das folhas, diferenciando folhas largas de folhas estreitas, e as touceiras, que frequentemente quebram a lógica de plantio implementada nas áreas agrícolas. Essas touceiras, com seu crescimento denso e irregular, são facilmente identificáveis, pois alteram o padrão de plantio e a distribuição das culturas. Esse procedimento foi crucial para o treinamento das redes neurais, garantindo que o modelo aprendesse a distinguir corretamente entre as classes de interesse.

No cálculo das métricas de desempenho, foi utilizada uma parte do conjunto de dados anotados. O conjunto foi dividido em duas partes: uma para treinamento da rede e modelos. A parte destinada ao treinamento foi usada para ensinar a rede a identificar as características do fundo e das daninhas nas imagens, enquanto a parte de validação foi utilizada para avaliar o desempenho da rede. As métricas de desempenho calculadas incluem a acurácia, que mede a proporção de previsões corretas em relação ao total de previsões feitas; o IoU (Intersection over Union), que quantifica a sobreposição entre a área predita e a área real das daninhas em uma imagem; o Dice, que é uma métrica de similaridade muito utilizada em segmentação, avalia a precisão da sobreposição entre as áreas preditas e reais; e o Loss, que mede a diferença entre as predições do modelo e os valores reais, sendo uma função essencial para otimizar o treinamento da rede. Essas métricas são fundamentais para avaliar a eficácia do modelo na tarefa de detecção de ervas daninhas, permitindo ajustar os parâmetros do modelo e melhorar sua performance [5][7].

A análise dos artigos incluiu a leitura dos resumos para uma seleção inicial, seguida de uma leitura mais detalhada dos textos completos para identificar as técnicas de CNNs aplicadas, os tipos de imagens de drones utilizadas e os resultados obtidos. Esse levantamento visou compreender o estado da arte, identificar as tendências mais recentes e possíveis áreas de pesquisa ainda não exploradas, fornecendo uma base sólida para o desenvolvimento do trabalho acadêmico em questão.

Adicionalmente, foi realizado um estudo sobre o balanceamento de classes (fundo e daninhas) no conjunto de amostras, com dimensões de 1024x1024 pixels. A análise considerou a contabilização dos pixels com valor 0, representando o fundo, e valor 1, representando as daninhas, a fim de avaliar a proporção de cada classe. Esse procedimento é essencial para entender como o balanceamento de classes pode impactar a eficácia das redes neurais, garantindo que a rede seja capaz de distinguir corretamente entre o fundo e as ervas daninhas, evitando viés na detecção [5][7].

3. RESULTADOS

A avaliação do desempenho dos modelos de detecção de ervas daninhas utilizou quatro métricas principais (Tabela 1), acurácia, Intersection over Union (IoU), coeficiente de Dice e função de perda (loss). A acurácia mede a proporção de previsões corretas em relação ao total de previsões feitas. O IoU avalia a precisão da segmentação ao medir a sobreposição entre a área segmentada e a área real das ervas daninhas. O coeficiente de Dice quantifica a similaridade entre a área segmentada e a área real, seguindo um padrão semelhante. A função de perda mede o erro do modelo durante o treinamento.

Na Tabela 1, são apresentados os resultados das métricas de acurácia obtidas pelas redes neurais treinadas para a identificação de plantas daninhas. Esta tabela destaca como cada modelo conseguiu distinguir entre plantas daninhas e o fundo presente nas imagens, evidenciando a capacidade do modelo em realizar uma classificação correta. Observa-se que a acurácia é uma métrica global que avalia a proporção de previsões corretas em relação ao total de amostras. Contudo, é importante ressaltar que, devido ao potencial desequilíbrio de classes — onde o fundo pode predominar em relação às áreas ocupadas por plantas daninhas — a acurácia, isoladamente, pode não refletir plenamente a eficiência da rede neural. Ainda assim, a alta acurácia observada indica uma robustez na performance do [illegible]mbora seja necessário considerar outras

métricas para uma análise mais abrangente.

Modelo	Acurácia	IoU	Dice	Loss
AttentionUnet	0.61	0.53	0.065	0.81
PSPNet	0.52	0.33	0.18	0.75
DeepLabv3+	0.77	0.25	0.36	0.78
ResUnet	0.92	0.86	0.92	0.31
SegFormer	0.95	0.93	0.95	0.08

Tabela 1. Métricas de acurácia da rede neural em relação à identificação de plantas daninhas

Na Tabela 2, o desempenho dos modelos de redes neurais no monitoramento de plantas daninhas foi mensurado através da análise de métricas em função do número de épocas (epochs) e do tamanho do lote (batch size) utilizado no treinamento. A variação desses hiperparâmetros influencia diretamente na capacidade dos modelos em aprender e generalizar a partir dos dados de treinamento. O número de épocas representa quantas vezes o modelo passou por todo o conjunto de dados de treinamento, enquanto o tamanho do lote define quantas amostras são processadas antes de atualizar os pesos do modelo. Um maior número de épocas, aliado a um tamanho de lote adequado, pode melhorar a performance do modelo ao garantir uma maior estabilidade no processo de aprendizado.

Na análise dos resultados, observou-se que modelos treinados com um número adequado de épocas e um tamanho de lote otimizado apresentaram melhor desempenho, com métricas mais altas de IoU e coeficiente de Dice, e uma redução significativa na função de perda ao longo do tempo. Isso indica que o modelo está se ajustando de forma eficaz às características das plantas daninhas, resultando em uma segmentação mais precisa e consistente.

Modelo	1° epoch	Outras epochs	Batch size
AttentionUnet	4h 30m	35m	4
PSPNet	3h 30m	15m	4
DeepLabv3+	2h 30m	16m	14
ResUnet	1h 15m	45m	4
SegFormer	4h 35m	1m 15s	4

Tabela 2. Desempenho dos modelos para monitoramento de plantas daninhas

4. DISCUSSÃO

Mediante as análises realizadas, a comparação de soluções que empregam Inteligência Artificial (IA), visão computacional (CV) ou uma combinação de ambas, revela duas principais problemáticas: os artigos que utilizaram IA frequentemente trabalharam com poucos mosaicos e, consequentemente, com um número reduzido de imagens para teste, enquanto os que adotaram CV usaram dados uniformes que não refletem a complexidade da solução necessária (AGARAP, 2017). Além disso, constatou-se que as principais técnicas de processamento baseadas em visão computacional, como a Transformada de Hough, Transformada Aleatória e Operações Morfológicas, apresentam limitações e têm dificuldades em escalar para considerável pós-processamento e têm baixa escalabilidade [3].

Diante desse cenário, optou-se por adotar soluções baseadas em IA, mais especificamente em Deep Learning. Considerando a necessidade de incluir mosaicos amplamente variados e garantir que a CNN produza resultados suficientemente generalizados para lidar com essa diversidade, é imperativo realizar testes com diversas redes neurais que representem o estado da arte no processamento digital de imagens. Logo, foram implementadas técnicas para prevenir overfitting, incluindo Data Augmentation, Dropout e regularização L2. Essas técnicas são essenciais para reduzir o risco de a rede aprender excessivamente as características dos dados de treinamento, o que poderia comprometer sua capacidade de generalização para novos dados de teste e validação [5].

5. CONCLUSÕES

O presente estudo evidencia a importância da detecção eficiente de ervas daninhas na cultura de cana-de-açúcar para maximizar a produtividade agrícola e promover a sustentabilidade. A análise comparativa entre diferentes abordagens baseadas em Inteligência Artificial (IA) e Visão Computacional (CV) revelou as limitações das técnicas tradicionais de processamento de imagens e destacou o potencial das Redes Neurais Convolucionais (CNNs) em lidar com a complexidade e a diversidade dos dados.

A análise comparativa das métricas de desempenho dos modelos revelou que a arquitetura Segformer se destacou como a mais eficaz, com a maior acurácia (0.95), Intersection over Union (IoU) (0.93), coeficiente de Dice (0.95) e a menor função de perda (0.08). Esses resultados evidenciam a capacidade superior do Segformer para lidar com a variabilidade e complexidade das imagens de drones, oferecendo uma solução robusta e confiável para a detecção de ervas daninhas. Além disso, a eficiência do Segformer em termos de tempo de treinamento e desempenho geral destaca a importância da escolha adequada da arquitetura de rede para aplicações práticas na agricultura. Dessa forma, o estudo propõe a aplicação da arquitetura Segformer como uma solução robusta para o monitoramento e controle de ervas daninhas na cultura de cana-de-açúcar, com vistas a maximizar a produtividade e minimizar os impactos ambientais.

Em relação aos resultados obtidos, o método de detecção utilizado neste estudo revelou a importância da detecção de ervas daninhas na agricultura de precisão. O modelo Segformer se destacou em termos de desempenho, sendo recomendado como a melhor solução para essa tarefa. No entanto, é importante destacar que esses resultados foram obtidos apenas em imagens multiespectrais, e os experimentos não foram realizados com imagens RGB. Fica em aberto a possibilidade de experimentos futuros com imagens RGB, a fim de investigar se os resultados se mantêm consistentes em diferentes tipos de imagens adquiridas por sensores embarcados em drones. A diferença de características entre imagens multiespectrais e RGB pode impactar a performance dos modelos, uma vez que as imagens multiespectrais oferecem mais informações sobre a reflectância das superfícies, o que pode ser benéfico na detecção de plantas daninhas.

8. REFERÊNCIAS

[1] AGARAP, F. An architecture combining convolutional neural network (CNN) and support vector machine (SVM) for image classification. 2017.

[2] ALDRICH, R. J. Prevendo reduções de rendimento de culturas por ervas daninhas. Weed Technology, v. 1, n. 3, p. 199-206, 1987. Disponível em: http://dx.doi.org/10.1017/S0890037X00029535. Acesso em: 20 ago. 2024.

[3] ARNAUD, F.; EVELINA, L.; FABRIZIO, R. Vision Inspection with Neural Networks. 2018. Dipartimento di Matematica e Informatica, University of Ferrara.

[4] GAZZIERO, D. L. P. Impactos econômicos da resistência de plantas daninhas a herbicidas no Brasil. Planta Daninha, v. 33, n. 3,
p. 1-7, 2015. DOI:10.1590/S0100-83582015000100015.

[5] LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. Nature, v. 521, p. 436-444, 2015. DOI: 10.1038/nature14539.

[6] MAPA. Ministério da Agricultura, Pecuária e Abastecimento. A cana-de-açúcar no Brasil: Desempenho e perspectivas. Brasília: MAPA, 2021.
[7] OLIVEIRA, J. F.; SANTOS, R. P. Sustentabilidade e o papel do etanol na matriz energética brasileira. Revista Brasileira de Energias Renováveis, v. 9, n. 3, p. 45-59, 2020. DOI: 10.1590/0104-876420204593.

[8] UNICA. União da Indústria de Cana-de-Açúcar. Produção e mercado. São Paulo, 2023. Disponível em: https://www.unica.com.br/producao-e-mercado/. Acesso em: 15 ago. 2024.